

Ayush Maheshwari

Senior Solutions Architect, NVIDIA

+91 9867216526 • ayush.hakmn@gmail.com
ayushbits.github.io • in ayushx • Ayush • ayushbits

Experience

NVIDIA

Senior Solutions Architect, NVAITC

Gurugram

Oct 2024 – Present

Vizzhy Inc

Lead Research Scientist, Language Model Team

Bengaluru

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

Research Intern, Document Cloud Team

Bengaluru

May – Aug 2021

NCETIS, IIT Bombay

Project Engineer

Mumbai

May 2016 – Dec 2018

Water & Climate Lab, IIT Gandhinagar

Junior Research Fellow

Gandhinagar

May 2015 – Apr 2016

Tata Consultancy Services

Systems Engineer

Mumbai

Oct 2011 – June 2013

Education

Indian Institute of Technology Bombay

Degree: Ph.D. in Computer Science and Engineering

Defence: July, 2024

Mumbai, India

Jan 2019 – Aug 2023

Thesis: Knowledge Integration in NLP Models using Constraint Ingestion and Generation

Advisor: Prof. Ganesh Ramakrishnan

Tata Institute of Social Sciences

Degree: M.Sc.

Mumbai, India

2013 – 2015

University College of Engineering

B.Tech. (CSE)

Rajasthan Technical University Kota, India

2007 – 2011

Open-Source contributions

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

[GitHub](#)

[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

- 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
- 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

Publications

1. DictDis: Dictionary Constrained Disambiguation for Improved NMT [↗](#)
Findings of EMNLP, 2024
Maheshwari, A., Jyothi P., Ramakrishnan G.
2. 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.
3. Data Programming using Semi-Supervision and Subset Selection [↗](#)
Findings of ACL, 2021
Maheshwari, A., Chatterjee O., Killamsetty K., Iyer R., Ramakrishnan G.
4. Joint Learning of Hyperbolic Label Embeddings for Hierarchical Multi-label Classification [↗](#)
EACL 2021
Maheshwari A., Chatterjee S., Ramakrishnan, G. Jagaralpudi S.

5. 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.
6. Adaptive mixing of auxiliary losses in supervised learning
AAAI, 2023 (Oral Presentation)
Sivasubramanian D., **Maheshwari, A.**, Prathosh AP, Shenoy P., Ramakrishnan G.
7. FAIR: Filtering of Automatically Induced Rules
EACL 2024
Bajpai, D.J., **Maheshwari, A.**, Hanawal, M. and Ramakrishnan, G.
8. Samayik: A Benchmark and Dataset for English-Sanskrit Translation
LREC-COLING 2024
Maheshwari, A., Gupta, A., Krishna, A., Singh, A.K., Ramakrishnan, G., Gourishetty, A.K. and Singla, J.
9. A Benchmark and Dataset for Post-OCR text correction in Sanskrit
Findings of EMNLP, 2022
Maheshwari, A., Singh, N., Krishna, A. and Ramakrishnan, G.,
10. Eigen: Expert-Informed Joint Learning Aggregation for High-Fidelity Information Extraction from Document Images
Machine Learning for Health (ML4H) at NeurIPS 2023
Singh, A., Venkatapathy, **Maheshwari, A.**, Narayan, P., Shetty, D.P. and Ramakrishnan, G.
11. OCR On-the-Go: Robust End-to-end Systems for Reading License Plates & Street Signs
International Conference on Document Analysis and Recognition, 2019
Saluja, R., **Maheshwari, A.**, Ramakrishnan, G., Chaudhuri, P. and Carman, M.
12. DynGAN: Generative Adversarial Networks for Dynamic Network Embedding
Graph Representation Learning Workshop at NeurIPS 2019
Maheshwari, A., Goyal, A., Hanawal, M.K. and Ramakrishnan, G.
13. Semi-supervised data programming in python
EMNLP 2022 - System Demonstrations
Abhishek, G., Ingole, H., Laturia, P., Dorna, V., **Maheshwari, A.**, Ramakrishnan, G. and Iyer, R.,
14. Rule Augmented Unsupervised Constituency Parsing
Findings of ACL, 2021
Sahay, A., Nasery, A., **Maheshwari, A.**, Ramakrishnan, G. and Iyer, R.
15. Representation learning on graphs by integrating content and structure information
11th International Conference on Communication Systems & Networks (COMSNETS)
Maheshwari, A., Goyal, A., Kumar, A., Hanawal, M.K. and Ramakrishnan, G.
16. Enhancing Low-Resource NMT with a Multilingual Encoder and Knowledge Distillation: A Case Study
LoResMT Workshop at ACL 2024
Roy, A., Ray, P., **Maheshwari, A.**, Sarkar, S. and Goyal, P.
17. Unsupervised learning of explainable parse trees for improved generalisation
IJCNN 2021

Sahay, A., **Maheshwari, A.**, Kumar, R., Ramakrishnan, G., Hanawal, M.K. and Arya, K.

18. Tale of tails using rule augmented sequence labeling for event extraction [!\[\]\(a22ba4e13c745edbf29e51af246c4c12_img.jpg\)](#)

Star-AI Workshop at AAAI 2019

Maheshwari, A., Patel, H., Rathod, N., Kumar, R., Ramakrishnan, G. and Bhattacharyya, P.

19. Entity Resolution and Location Disambiguation in the Ancient Hindu Temples Domain using Web Data [!\[\]\(cdf2842d82858164c68c92720a337fb9_img.jpg\)](#)

NAACL 2018 - System Demonstrations

Maheshwari, A., Kumar, V., Ramakrishnan, G. and Nath, J.S.