

Ayush Maheshwari

PhD Candidate (Final Year), IIT Bombay (India)

✉ ayusham@cse.iitb.ac.in • 🌐 cse.iitb.ac.in/~ayusham • in ayushx
📷 Ayush • 🐙 ayushbits

Education

Indian Institute of Technology Bombay

Ph.D. (CSE), Recipient of Ekal Fellowship

Mumbai, India

2019–Present

Advisors: Prof. Ganesh Ramakrishnan & Prof. Manjesh K. Hanawal

Tata Institute of Social Sciences

M.Sc.

Mumbai, India

2013–2015

University College of Engineering

B.Tech. (CSE)

Rajasthan Technical University Kota, India

2007–2011

Research Interests

- Develop machine learning based methods that generalize to new domains with limited resources.
- Lexically constrained neural machine translation
- Data programming, semi-supervised learning, subset selection

Experience

Adobe Research

Research Intern, Document Cloud Team

Bengaluru

May 2021 – Aug 2021

NCETIS, IIT Bombay

Project Engineer

Mumbai

2016–2018

Water & Climate Lab, IIT Gandhinagar

Junior Research Fellow

2015–2016

Tata Consultancy Services

Systems Engineer

Mumbai

2011–2013

Open-Source contributions

- **SPEAR** : Data-Programming python library to create training data. [Repo link](#)
- **UDAAN** : Post-editing workbench to quickly create publishable translations. [Paper link](#)

Computer Skills

Programming Languages: Python, C, C++, R, PHP, Javascript/jQuery, shell script

Libraries: PyTorch, NumPy, scikit-learn

Applications: \LaTeX , Google Cloud Platform

Publications

1. **Maheshwari, A.**, Singh N., Krishna A., Ramakrishnan G. **A Benchmark and Dataset for Post-OCR text correction in Sanskrit.** Findings of EMNLP, 2022
2. **Maheshwari, A.**, Abhishek G., Ingole H., Laturia P., Dorna V., Ramakrishnan G., Iyer R., **SPEAR: Semi-supervised Data Programming in Python.** EMNLP Demonstrations, 2022
3. **Maheshwari, A.**, Killamsetty K., Ramakrishnan G., Iyer R., Danilevsky M., & Popa L. **Learning to Robustly Aggregate Labeling Functions for Semi-supervised Data Programming.** Findings of ACL, 2022
4. **Maheshwari, A.**, Chatterjee O., Killamsetty K., Iyer R., & Ramakrishnan G. **Data Programming using Semi-Supervision and Subset Selection.** Findings of ACL, 2021
5. Sahay A., Nasery A., **Maheshwari A.**, Ramakrishnan G. and Iyer R. **Rule Augmented Unsupervised Constituency Parsing.** Findings of ACL, 2021
6. Sahay A, **Maheshwari A.**, Kumar R., Ramakrishnan, G., Hanawal M. **Unsupervised Learning of Explainable Parse Trees for Improved Generalisation.** IJCNN 2021
7. Chatterjee S.*, **Maheshwari A.***, Ramakrishnan, G. & Jagaralpudi S. **Joint Learning of Hyperbolic Label Embeddings for Hierarchical Multi-label Classification.** To appear at European Chapter of the Association for Computational Linguistics (EACL) 2021
8. Saluja, R., **Maheshwari, A.**, Ramakrishnan, G., Chaudhuri, P., & Carman, M. **OCR On-the-Go: Robust End-to-end Systems for Reading License Plates & Street Signs.** In 2019 International Conference on Document Analysis and Recognition (ICDAR) (pp. 154–159).
9. **Maheshwari, A.**, Goyal, A., Kumar, A., Hanawal, M., & Ramakrishnan, G. **Representation Learning on Graphs by Integrating Content and Structure Information.** In 2019 11th International Conference on Communication Systems & Networks (COMSNETS) (pp. 88–94).
10. **Maheshwari, A.**, Patel, H., Rathod, N., Kumar, R., Ramakrishnan, G., & Bhattacharyya, P. **Tale of tails using rule augmented sequence labeling for event extraction** In 2020 Ninth International Workshop on Statistical Relational AI at the 34th AAAI Conference on Artificial Intelligence (AAAI)
11. **Maheshwari, A.**, Goyal, A., Hanawal, M., & Ramakrishnan, G. **DynGAN: Generative Adversarial Networks for Dynamic Network Embedding.** In Graph Representation Learning Workshop at NeurIPS 2019.
12. **Maheshwari, A.**, Kumar, V., Ramakrishnan, G., & Nath, J. **Entity Resolution and Location Disambiguation in the Ancient Hindu Temples Domain using Web Data.** In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Demonstrations (pp. 46–50).

Services

Teaching Assistantships

- Artificial Intelligence and Machine Learning Autumn 2019
- Organization of Web information Spring 2020
- Information Retrieval & Mining for Hypertext & the Web Autumn 2020
- Speech, Natural Language Processing and the Web Spring 2020
- Digital Health Informatics Autumn 2021

Conference Reviewing

- Association for Computational Linguistics (ACL) 2021
- Empirical Methods in Natural Language Processing (EMNLP) 2020-2022
- European Association for Computational Linguistics(EACL) 2020
- North-American Association for Computational Linguistics (NAACL) 2019, 2020
- Association for the Advancement of Artificial Intelligence Conference (AAAI) 2020