

Shruti Singh

Department of Computer Science & Engineering
Indian Institute of Technology, Gandhinagar
Gujarat, India

Email: singh_shruti@iitgn.ac.in
Links: [Webpage](#), [Github](#)
Phone: +91 95371 60842

Education

Indian Institute of Technology, Gandhinagar <i>Ph.D. Computer Science (Prime Minister's Research Fellow)</i> Advisor: Mayank Singh GPA: 9.07/10.0	2019 - 2024
Dhirubhai Ambani Institute of Information & Communication Technology <i>Bachelor (Hons.) in Information and Communication Technology with Minor in Computational Science</i> GPA: 8.25/10.0	2013 - 2017

Research Interests

Broad: Natural Language Processing, Deep Learning, Recommendation Systems

Teaching Experience

Teaching assistant for the following courses:

Natural Language Processing (Autumn 2020 - Head TA), Recommendation Systems (Spring 2022 - Head TA), Databases (Spring 2022, Spring 2021, Spring 2020), Introduction to Python (Spring 2019, Autumn 2021)

Coursework

Natural Language Processing, Machine Learning, Deep Learning, Advanced Machine Learning, Artificial Intelligence, Information Retrieval, Graduate Algorithms, Nature Inspired Computing, Learning With Graphs, Brain Imaging Methods, Computational Neuroscience, Object Oriented Programming, Data Structures, High Performance Computing, GPU Programming, Data Analysis and Visualization

Research and Industry Experience

LINGO - IIT Gandhinagar

- Representation Learning – My PhD research focuses on learning representations of long documents (specifically scientific articles) that are task-independent, interpretable, and efficiently encode knowledge from research papers.
- Multilingual Abusive Comment Detection – We pretrain multilingual language models to detect abusive comments in 10+ low-resource Indic languages on Moj dataset with an accuracy of 88.5%.
- Understanding Hindi-English Search in the Brain – We investigate the possibility of a cross-linguistically shared mental lexicon of Hindi and English in various conceptual associations in humans.

RAXTER.IO

Feb, 2018 - June, 2019

Research Engineer

- Implemented a Q-Learning based literature recommendation system that adapts to the evolving research needs of the user with time.
- Implemented a system for information extraction from scientific literature, that extracts the problem statement, approach, evaluation techniques, results, and datasets for construction of scientific literature graph.

Sprinklr Inc.

Jan, 2017 - Feb, 2018

Product Engineer

- Worked on near de-duplication of large-scale real-time incoming web data of the scale of 25 million documents/day, ensuring minimal latency.
- Implemented language detection module for close languages Malay and Indonesian.
- Setup the natural language pipeline for Asia Pacific languages on the Sprinklr Listening platform.

Rygbee Inc.

May, 2016 - July, 2016

Research Intern

- Developed a recommendation system for scientific literature and MOOCs. Developed an in-house NLP pipeline for pre-processing and recommendation, used Elasticsearch for indexing papers.

Technical Skills

Programming*Over 5000 lines:* Java, Python*Over 1000 lines:* C, C++**Technologies**

Elasticsearch, MongoDB, MySQL, Apache Kafka, Apache Lucene, Flask, Docker, Git, CUDA, OpenMP, StanfordCoreNLP, Spacy, Wordnet, NLTK, OpenCV, Gensim, Sklearn, Surprise, Pytorch

Awards

- October 2021: Selected for the Prime Minister's Research Fellow (PMRF) Fellowship.
- July 2021: Selected for Science of Science Summer School 2021 hosted by Syracuse University, among 57 out of 289 applicants.
- Aug 2020: Selected for DLRL Summer School 2020 hosted by CIFAR and Mila, among 250 out of 1000+ applicants.
- 2013: Certificate of Merit for being in top 1.5% students in the CBSE AISSCE 2013.
- 2011: Dhirubhai Ambani SSC Merit Reward for outstanding performance at CBSE AISSCE 2011.

Responsibilities

- Student Volunteer for NAACL 2021, ACL-IJCNLP 2021, IndoML 2020, ACL 2020, ICLR 2020, ACM India Annual Event 2019, and the Inter-Research-Institute Student Seminar (IRISS) 2019.

- 2015-2016: Editor Entelechy - DA-IICT University Magazine.
- 2015-2016: Core Committee Member - Programming Club DA-IICT.

Publications

The Inefficiency of Language Models in Scholarly Retrieval: An Experimental Walk-through

Findings of ACL 2022

Shruti Singh, Mayank Singh

TweeNLP: A Twitter Exploration Portal for Natural Language Processing

Demonstrations, ACL-IJCNLP 2021

Viraj Shah, Shruti Singh, Mayank Singh

[Portal](#), [Paper](#)

COMPARE: A Taxonomy and Dataset of Comparison Discussions in Peer Reviews

Joint Conference on Digital Libraries 2021

Shruti Singh, Mayank Singh, Pawan Goyal

[Paper](#)

Genealogy Tree Construction of Research Papers

Young Researcher Symposium, CODS-COMAD 2021

Mihir Jain*, Tanmay Sharma*, Aman*, Shruti Singh, Mayank Singh

[Extended Abstract](#)

Understanding Attention: In Minds and Machines

ML-RSA at NeurIPS 2020

Shriraj Sawant*, Shruti Singh*

[Paper](#)

NLPExplorer: Exploring the Universe of NLP Papers

Demonstrations, European Conference on Information Retrieval 2020

Monarch Parmar*, Naman Jain*, Pranjali Jain*, Jayakrishna Sahit*, Soham Pachpande*, Shruti Singh, and Mayank Singh

[Portal](#), [Paper](#)