Shruti Singh

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

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

Education

Indian Institute of Technology, Gandhinagar   2019 - 2024
Ph.D. Computer Science (Prime Minister's Research Fellow)
Advisor: Mayank Singh
GPA: 9.2/10.0

Dhirubhai Ambani Institute of Information & Communication Technology   2013 - 2017
Bachelor (Hons.) in Information and Communication Technology with Minor in Computational Science
GPA: 8.25/10.0

Central Board of Secondary Education
Higher Secondary School - Kendriya Vidyalaya   2013
Aggregate: 95.6%
Senior Secondary School - Kendriya Vidyalaya   2011
GPA: 10/10

Research Interests

Research Goal: My long-term goal is to develop tools that can assist researchers and democratize the entry of marginalized communities into research.
Broad: Natural Language Processing, Deep Learning, Recommender Systems, Cognitive Linguistics
Specific: NLP for Scientific Literature

Fellowships and Grants

- May 2022: Selected for the ORKG Curation Grant 2022 by TIB Hannover, Germany.
- October 2021: Selected as a Prime Minister's Research Fellow (PMRF).

Awards

- July 2022: Selected for DLRL Summer School 2022 hosted by CIFAR and Mila, selected among 15 students to present my research idea to experts in the field.
- 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.

Publications

The Inefficiency of Language Models in Scholarly Retrieval: An Experimental Walkthrough
Findings of ACL 2022   Paper
Shruti Singh, Mayank Singh

COMPARE: A Taxonomy and Dataset of Comparison Discussions in Peer Reviews
Joint Conference on Digital Libraries 2021   Paper
Shruti Singh, Mayank Singh, Pawan Goyal

TweeNLP: A Twitter Exploration Portal for Natural Language Processing
Demonstrations, ACL-IJCNLP 2021   Portal, Paper
Viraj Shah, Shruti Singh, Mayank Singh

Genealogy Tree Construction of Research Papers
Young Researcher Symposium, CODS-COMAD 2021   Extended Abstract
Mihir Jain*, Tanmay Sharma*, Aman*, Shruti Singh, Mayank Singh

Understanding Attention: In Minds and Machines   Paper
ML-RSA at NeurIPS 2020
Shriraj Sawant*, Shruti Singh*

NLPExplorer: Exploring the Universe of NLP Papers   Portal, Paper
Demonstrations, European Conference on Information Retrieval 2020
Monarch Parmar*, Naman Jain*, Pranjali Jain*, Jayakrishna Sahit*, Soham Pachpande*,
Shruti Singh, and Mayank Singh

Research and Industry Experience

LINGO - IIT Gandhinagar   Aug, 2019 -
- 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.

Bosch Research   Aug, 2022 - Dec, 2022
Research Intern
Representation learning for procedural documents by representing long documents as graphs.

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.

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

Service

- Student Volunteer for ACM-IKDD Summer School 2022 on Data Science, NAACL 2021, ACL-IJCNLP 2021, IndoML 2020, ACL 2020, ICLR 2020, ACM-W India Workshop 2020, ACM India Annual Event 2020, and the Inter-Research-Institute Student Seminar (IRISS) 2020.
- Webmaster for IndoML 2020.
- 2015-2016: Editor Entelechy - DA-IICT University Magazine.
- 2015-2016: Core Committee Member - Programming Club DA-IICT.

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