

# Yash Kumar Atri, Ph.D.

Postdoc at University of Virginia

1919 ivy road, University of Virginia, USA

[yashkumaratri@gmail.com](mailto:yashkumaratri@gmail.com)

[yashkumaratri.github.io](https://yashkumaratri.github.io)



## ACADEMIC APPOINTMENTS

### University of Virginia, School of Data Science

Postdoctoral Research Associate

USA

July 2024 - Now

- **Advisor:** Prof. Tom Hartvigsen
- **Topic:** Controlling Hallucinations and Improving Faithfulness in Large Language Models through Model Editing

## EDUCATION

### Indraprastha Institute of Information Technology (IIIT-Delhi)

Doctor of Philosophy (Ph.D.)

Delhi, India

2020 - 2024

- **Advisors:** Prof. Vikram Goyal, Prof. Tanmoy Chakraborty
- **Thesis Title:** Advancing text summarization with conscience, comprehension, and multimodality

### Jaypee University Anoopshahr

Bachelor of Technology (B.Tech.) in Computer Science and Engineering

Anoopshahr, India

2014 - 2018

- **Advisors:** Dr. Amit Kumar
- **Thesis Title:** Machine translation in Indian languages

## PUBLICATIONS

C=CONFERENCE, J=JOURNAL, P=PREPRINT

- C.5** Yash Kumar Atri, Arun Iyer, Tanmoy Chakraborty, Vikram Goyal, [Promoting Topic Coherence and Inter-Document Consorts in Multi-Document Summarization via Simplicial Complex and Sheaf Graph](#), Conference on Empirical Methods in Natural Language Processing (EMNLP) 2023
- J.2** Yash Kumar Atri, Vikram Goyal, Tanmoy Chakraborty, [Multi-Document Summarization using Selective Attention Span and Reinforcement Learning](#), IEEE Transactions on Audio, Speech, and Language Processing (TASL) 2023
- C.4** Yash Kumar Atri, Vikram Goyal, Tanmoy Chakraborty, [Fusing Multimodal Signals on Hyper-complex Space for Extreme Abstractive Text Summarization \(TL;DR\) of Scientific Contents](#), ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD) 2023
- P.1** Yash Kumar Atri, Vikram Goyal, Tanmoy Chakraborty, [Exploiting Representation Bias for Data Distillation in Abstractive Text Summarization](#), arXiv Preprint (Arxiv) 2023
- C.3** Priyanshi Gupta, Yash Kumar Atri, Apurva Nagvenkar, Sourish Dasgupta, Tanmoy Chakraborty, [Inline Citation Classification using Peripheral Context and Time-evolving Augmentation](#), Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) 2023
- C.2** Vaibhav Pulastya, Gaurav Nuti, Yash Kumar Atri, Tanmoy Chakraborty, [Assessing the Quality of the Datasets by Identifying Mislabeled Samples](#), IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) 2021
- J.1** Yash Kumar Atri\*, Shraman Pramanick\*, Vikram Goyal, Tanmoy Chakraborty, [See, Hear, Read: Leveraging Multimodality with Guided Attention for Abstractive Text Summarization](#), Elsevier Knowledge-Based Systems (KBS) 2021
- C.1** Alvin Dey\*, Tanya Chowdhury\*, Yash Kumar Atri\*, Tanmoy Chakraborty, [Corpora Evaluation and System Bias Detection in Multi-Document Summarization](#), Findings of the Association for Computational Linguistics at the Conference on Empirical Methods in Natural Language Processing (EMNLP Findings) 2020

## INVITED TALKS

### ACM-India's Symposium for Academic Research and Careers for Students

Lightning Talk and Poster Presentation

NISER Bhubhaneshwar, India

2023

### International Conference on Natural Language Processing (ICON)

Tutorial on Building Blocks of AI-Driven Mental Health Counseling

University of Goa, India

2023

### RIISE 2024

Poster Presentation

IIIT Delhi, India

2023

## AWARDS AND ACHIEVEMENTS

- **DAAD AInet Fellow** 2024  
Postdoc-NeT-AI, Germany
- **Microsoft Travel Grant, iHub-Anubhuti Travel Grant** 2023  
For Travel to Singapore for EMNLP
- **Microsoft Travel Grant, iHub-Anubhuti Travel Grant, ACM-India Travel Grant (IARCS)** 2023  
For Travel to Long Beach, USA for KDD

## INDUSTRY EXPERIENCE

### Laboratory for Computational Social Systems (LCS2), IIIT Delhi

Research Associate

Delhi, India

Sep. 2019 - Dec. 2019

### Lumiq.ai

Software Engineer (Data Science)

Noida, India

Jul. 2018 - Aug. 2019

### Lumiq.ai

AI Engineer (Intern)

Noida, India

Feb. 2018 - Jun. 2018

### Primelite Technologies

Data Scientist (Intern)

Noida, India

May 2017 - Jul. 2017

### Trignosoft Technologies

Android Developer (Virtual Intern)

Noida, India

May 2017 - Jun. 2017

## PROJECTS

- **Aerial Cactus Identification:** Developed a high-accuracy model leveraging VGG16 architecture to identify cactus species in aerial images, achieving a top accuracy of 99.98% on the Kaggle platform.
- **Predict Future Sales:** Engineered robust feature extraction techniques for an e-commerce sales prediction model, achieving a root mean square error (RMSE) of 0.9239, optimizing predictive accuracy on structured datasets from Kaggle.
- **Histopathologic Cancer Detection:** Applied a ResNet101 model to classify cancerous tissue in pathology image patches, reaching 96.69% classification accuracy through augmentation techniques and extensive fine-tuning.
- **TalkingData Ad Tracking Fraud Detection:** Crafted targeted feature engineering and optimized detection algorithms to identify fraudulent activity in financial datasets with 97.87% accuracy, addressing complex data challenges on the Kaggle platform.
- **Jigsaw Unintended Bias in Toxicity Classification:** Implemented an LSTM-based architecture with specialized embeddings to filter toxic comments from online platforms, achieving 93% precision in toxicity filtering across diverse datasets.
- **Machine Translation in Indian Languages:** Designed and developed an encoder-decoder model with Hierarchical Attention for Hindi-English machine translation, achieving a significant boost in translation accuracy by handling complex linguistic nuances.
- **Detection and Interpretation of English Puns:** Scored 81% accuracy in SEMEVAL 2017 Task 7 by categorizing sentences with or without puns, employing a parallel network structure using CNN, GRU, and ANN models to handle language subtleties effectively.

## TAS AND COMMUNITY WORK

### Reviewer

ICLR24, ARR24, EMNLP23, ACL23, BDA23, ICON22, EMNLP22, ASONAM22

-

-

### Organizer & Coordinator

BDA 2023 Conference, ICON 2023 Conference, ACSS 2022, 2021, 2020, Cofad 2020 Workshops

IIIT Delhi, India

2020-2023

### Teaching Assistant

CSE557 (Winter 2020), CSE506 (Monsoon 2020), CSE557 (Winter 2021), CSE506 (Monsoon 2021)

IIIT Delhi, India

2020-2022

## REFERENCES

### 1. Prof Tom Hartvigsen

Assistant Professor, School of Data Science

University of Virginia

Email: hartvigsen@virginia.edu

Relationship: Postdoc Advisor

### 2. Prof Vikram Goyal

Professor, Dept. of Computer Science

Indraprastha Institute of Technology Delhi (IIIT Delhi), India

Email: vikram@iiitd.ac.in

Relationship: PhD Advisor

### 3. Prof Tanmoy Chakraborty

Associate Professor, Dept. of Electrical Engineering

Indian Institute of Technology Delhi (IIT Delhi), India

Email: tanchak@iitd.ac.in, chak.tanmoy.iit@gmail.com\*

Relationship: PhD Advisor