

# Rajdeep Mukherjee

PH.D. RESEARCH SCHOLAR (INSTITUTE), CNERG, DEPT. OF CSE, IIT KHARAGPUR

P-107/A, Unique Park, Behala, Kolkata - 700034, West Bengal, India

+91-9836205250 | [rajdeep1989.iitkgp@gmail.com](mailto:rajdeep1989.iitkgp@gmail.com) | [rajdeep345.github.io](https://github.com/rajdeep345) | [Rajdeep Mukherjee](#) | [rajdeep345](#) | [rajdeepMukherjee89/](#)

## Summary

I am working as a Senior AI Research Scientist with NielsenIQ. Also, I am a final-year Ph.D. Candidate (thesis submitted) at the Dept. of CSE, IIT Kharagpur advised by [Prof. Pawan Goyal](#) and [Prof. Niloy Ganguly](#). My primary area of research is Natural Language Processing (NLP). My Ph.D. thesis focuses on addressing the modeling and evaluation requirements of text summarization solutions across multiple domains such as e-commerce, finance, disaster mitigation, law, and tourism. My research interests include Large Language Models (LLMs), Numeracy, Factuality, Instruction Tuning, Parameter-Efficient Fine-tuning, Extreme Multi-label Classification, Information Extraction, and Text Summarization. I will be happy to talk about ideas/collaborations/opportunities along these directions. I am also interested in the areas of AI for Health and Social Good.

## Education

### Indian Institute of Technology (IIT), Kharagpur

PHD IN COMPUTER SCIENCE AND ENGINEERING

Kharagpur, India

Jan 2018 - May 2023

- Thesis Topic - Domain-Specific Text Summarization: Modeling and Evaluation
- Advisors: Prof. Pawan Goyal, Prof. Niloy Ganguly

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

M.TECH IN COMPUTER SCIENCE AND ENGINEERING

Delhi, India

Jul 2015 - Jul 2017

- CGPA: 9.11, Ranked 1st in Information Security Specialization (with thesis)
- Supervisor: [Prof. Sambuddho Chakravarty](#) | [M.Tech Thesis](#)

### Jadavpur University, Kolkata

B.E. IN INFORMATION TECHNOLOGY

Kolkata, India

Jul 2007 - Jun 2011

- CGPA: 8.59, University Gold Medalist, TCS Best Student Project awardee.
- Supervisor: [Prof. Samiran Chattopadhyay](#) | [Bachelors Thesis](#)

### Kendriya Vidyalaya (ONGC) Chandkheda, Ahmedabad, Gujarat

ALL INDIA SENIOR SCHOOL CERTIFICATE EXAMINATION (AISSCE)

Ahmedabad, India

2007

- Percentage: 91.8 %

### Kendriya Vidyalaya (ONGC) Chandkheda, Ahmedabad, Gujarat

ALL INDIA SECONDARY SCHOOL EXAMINATION (AISSE)

Ahmedabad, India

2005

- Percentage: 94.2 %
- Adjudged School Topper; 2nd place, KVS Ahmedabad Region (with 47+ schools)

## Work Experience

### NielsenIQ

SENIOR AI RESEARCH SCIENTIST

Pune, India

Jan 2024 - Present

- Working on various problems related to automation of market research using AI/NLP.

### Amazon Science

APPLIED SCIENTIST II INTERN

Bangalore, India

Apr - Oct 2023

- Worked towards improving customer search experience on e-commerce websites using Large Language Models.
- Industry Track paper accepted at WSDM 2024, Full paper under review.

### IBM Research Labs

RESEARCH INTERN

Delhi, India

May - Jul 2018

- Worked towards designing a CNN-based solution for identifying and recognizing bird sounds.
- Secured **7th rank worldwide** (1st in India) in the **Bird Audio Detection** task as part of [IEEE AASP DCASE2018 Challenge](#).

### PricewaterhouseCoopers (PwC) India

TECHNOLOGY CONSULTANT

Kolkata, India

Aug 2011 - Jul 2015

- Worked as a Business Intelligence (BI) developer and Data Modeler in the Advisory Line of Service (LOS).
- Led projects in data warehousing, data modeling and BI using Microsoft technologies like MSBI suite, Visual Studio, MS SQL Server 2012 etc.

## Selected Publications

---

### **CONTRASTE: Supervised Contrastive Pre-training With Aspect-based Prompts For Aspect Sentiment Triplet Extraction**

**Rajdeep Mukherjee**, Nithish Kannen, Saurabh Kumar Pandey, Pawan Goyal

[Paper](#) | [Code](#) | [Poster](#) | [Citation](#)

EMNLP 2023

### **MILDSum: A Novel Benchmark Dataset for Multilingual Summarization of Indian Legal Case Judgments**

Debtanu Datta, Shubham Soni, **Rajdeep Mukherjee**, Saptarshi Ghosh

[Paper](#) | [Dataset & Code](#) | [Poster](#) | [Slides](#) | [Video](#) | [Citation](#)

EMNLP 2023

### **ECTSum: A New Benchmark Dataset For Bullet Point Summarization of Long Earnings Call Transcripts**

**Rajdeep Mukherjee**, Abhinav Bohra, Akash Banerjee, Soumya Sharma, Manjunath Hegde, Afreen Shaikh, Shivani Shrivastava, Koustuv Dasgupta, Niloy Ganguly, Saptarshi Ghosh, Pawan Goyal

[Paper](#) | [Dataset & Code](#) | [Poster](#) | [Slides](#) | [Video](#) | [Citation](#)

EMNLP 2022

### **MTLTS: A Multi-Task Framework To Obtain Trustworthy Summaries From Crisis-Related Microblogs**

**Rajdeep Mukherjee**, Uppada Vishnu\*, Hari Chandana Peruri\*, Sourangshu Bhattacharya, Koustav Rudra, Pawan Goyal, Niloy Ganguly

[Paper](#) | [Code](#) | [Poster](#) | [Slides](#) | [Video](#) | [Citation](#)

WSDM 2022

### **PASTE: A Tagging-Free Decoding Framework Using Pointer Networks for Aspect Sentiment Triplet Extraction**

**Rajdeep Mukherjee**, Tapas Nayak, Yash Butala, Sourangshu Bhattacharya, Pawan Goyal

[Paper](#) | [Code](#) | [Poster](#) | [Slides](#) | [Video](#) | [Citation](#)

EMNLP 2021

### **Understanding the Role of Affect Dimensions in Detecting Emotions from Tweets: A Multi-task Approach**

**Rajdeep Mukherjee**, Atharva Naik\*, Sriyash Poddar\*, Soham Dasgupta, Niloy Ganguly

[Paper](#) | [Code](#) | [Blog](#) | [Poster](#) | [Slides](#) | [Video](#) | [Citation](#)

SIGIR 2021

### **Reproducibility, Replicability and Beyond: Assessing Production Readiness of Aspect Based Sentiment Analysis in the Wild**

**Rajdeep Mukherjee\***, Shreyas Shetty\*, Subrata Chattopadhyay, Subhadeep Maji, Samik Datta, Pawan Goyal

[Paper](#) | [Code](#) | [Slides](#) | [Video](#) | [Citation](#)

ECIR 2021

### **Read what you need: Controllable Aspect-based Opinion Summarization of Tourist Reviews**

**Rajdeep Mukherjee**, Hari Chandana Peruri, Uppada Vishnu, Pawan Goyal, Sourangshu Bhattacharya, Niloy Ganguly

[Paper](#) | [Code](#) | [Slides](#) | [Video](#) | [Citation](#)

SIGIR 2020

## Professional Activities

---

### **Program Committee Member**

AAAI[’24,’23], WSDM’24, EMNLP[’23,’22], ACL’23, EACL’23, SIGIR’22

### **Reviewer/Sub-Reviewer for Conference**

NAACL[’24,’22], AAAI[’24,’23], WSDM’24, EMNLP[’23,’22,’20,’19], ACL[’23,’22], EACL’23, SIGIR[’22,’21,’20], NEURIPS’21, SDM’21, CODS-COMAD[’21,’19], LREC’20

### **Reviewer/Sub-Reviewer for Journal**

TWEB[’22,’20], COMPUTER SPEECH AND LANGUAGE’22, IJDRR’21, TCSS’20

## Teaching Assistance

---

Spring '22	<b>CS60075: Natural Language Processing</b> , Instructor: Prof. Pawan Goyal	IIT Kharagpur
Autumn '21	<b>CS60092: Information Retrieval</b> , Instructor: Prof. Saptarshi Ghosh	IIT Kharagpur
Autumn '21	<b>noc21-cs102: Natural Language Processing</b> , Instructor: Prof. Pawan Goyal	NPTEL
Spring '21	<b>CS60075: Natural Language Processing</b> , Instructor: Prof. Pawan Goyal	IIT Kharagpur
Autumn '20	<b>CS60083: Parameterized Algorithms</b> , Instructors: Prof. Palash Dey, Prof. Sudeshna Kolay	IIT Kharagpur
Spring '20	<b>CS60092: Information Retrieval</b> , Instructors: Prof. Pawan Goyal, Prof. Saptarshi Ghosh	IIT Kharagpur
Autumn '19	<b>CS29003: Algorithms-1 Lab</b> , Instructors: Prof. Pawan Goyal, Prof. Animesh Mukherjee	IIT Kharagpur
Autumn '19	<b>noc19-cs56: Natural Language Processing</b> , Instructor: Prof. Pawan Goyal	NPTEL
Spring '19	<b>Programming and Data Structures Lab</b> , Instructor: Prof. Rogers Mathew	IIT Kharagpur
Autumn '18	<b>Programming and Data Structures Lab</b> , Instructors: Prof. Rogers Mathew, Prof. Sudebkumar Prasant Pal	IIT Kharagpur
Spring '17	<b>CSE352/CSE552: Security Engineering</b> , Instructor: Prof. Sambuddho Chakravarty	IIIT Delhi

## Scholarships & Travel Grants

---

2022	<b>Microsoft Research Travel Grant</b> , awarded for attending WSDM 2022.	Microsoft Research
2022	<b>ACM-India and IARCS Travel Grant</b> , awarded for attending WSDM 2022.	ACM-India IARCS
2022	<b>SIGIR Student Travel Grant</b> , awarded for registering the paper.	WSDM 2022
2021	<b>SIGIR Student Travel Grant</b> , awarded for registering the paper and attending the conference (virtually).	SIGIR 2021
2020	<b>SIGIR Student Travel Grant</b> , awarded for registering the paper and attending the conference (virtually).	SIGIR 2020

## Honors & Awards

---

2013	<b>PwC Experience Award</b> , awarded twice for demonstrating PwC Core values and behaviors.	Kolkata, India
2012	<b>PwC Above and Beyond Award</b> , awarded by Sudipto Ghosh, Partner, PwC.	Kolkata, India
2011	<b>TCS Best Student Project Award and University Gold Medal</b> , jointly awarded by Tata Consultancy Services and Jadavpur University.	Kolkata, India
2010	<b>2nd place, IBM Technical Quiz</b> , INSTRUO 2010 held at IEST, Shibpur in the Technology Area on Green IT.	Kolkata, India
2005	<b>2nd place, AISSE, KVS Ahmedabad Region</b> , selected for CSIR Program on Youth for Leadership in Science.	Ahmedabad, India

## Extracurricular Activities

---

2021	<b>Judge</b> , Open IIT: Eastern Vocal Competition	Kharagpur, India
2018	<b>Permanent Membership</b> , Eastern Technology Music Society (ETMS), IIT Kharagpur.	Kharagpur, India
2015	<b>Permanent Membership</b> , AudioBytes, Music Club, IIIT Delhi.	Delhi, India
2013	<b>Silvers Plate Runners Up</b> , Merchants' Cup Badminton Tournament.	Kolkata, India
2011	<b>Multiple awards</b> , National / State level Singing Competitions.	Kolkata, India
2010	<b>Member, Organizing Committee</b> , SRIJAN 2010, Annual Technical Fest, Jadavpur University.	Kolkata, India

## References

---

### Dr. Niloy Ganguly

Professor, Department of Computer Science and Engineering, IIT Kharagpur

[NILOY@CSE.IITKGP.AC.IN](mailto:NILOY@CSE.IITKGP.AC.IN)

### Dr. Pawan Goyal

Associate Professor, Department of Computer Science and Engineering, IIT Kharagpur

[PAWANG@CSE.IITKGP.AC.IN](mailto:PAWANG@CSE.IITKGP.AC.IN)

### Dr. Saptarshi Ghosh

Associate Professor, Department of Computer Science and Engineering, IIT Kharagpur

[SOURANGSHU@CSE.IITKGP.AC.IN](mailto:SOURANGSHU@CSE.IITKGP.AC.IN)