Suman Adhya

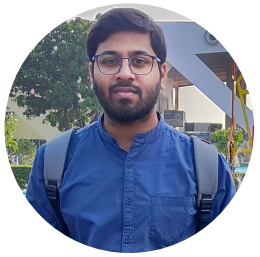
PhD Scholar at SMCS in IACS

adhyasuman.github.io

adhyasuman30@gmail.com

linkedin.com/in/sumanadhya AdhyaSuman Google Scholar

IACS, Jadavpur, Kolkata-700032, West Bengal, India



Vision Statement

My doctoral research advanced the **robustness**, **interpretability**, and **efficiency** of neural topic models by improving **representation learning**, preventing posterior collapse, and applying **knowledge distillation**. Building on this foundation, my current agenda targets two fronts for LLMs: **mechanistic interpretability**—uncovering how internal computations yield behavior—and **algorithmic scaling**—making large models more tractable and sustainable. My ultimate goal is to leverage my skills in representation learning and algorithmic design to help build LLMs that are **auditable, trustworthy, and resource-efficient**.

Short Bio

I am a Ph.D. candidate at the School of Mathematical & Computational Sciences (SMCS) in IACS, with a specialization in topic modeling for NLP. My research focuses on enhancing the quality of topic representations, improving interpretability, and increasing the efficiency of topic models, with publications in venues such as ACL 2025, NAACL 2024, EACL 2023, ECIR 2023, and IEEE Transactions on Artificial Intelligence. I will submit my doctoral thesis for review in August 2025 and am actively seeking a full-time research position in either industry or academia.

Education

Ongoing **Oct 2025**	**Indian Association for the Cultivation of Science** Ph.D. from the School of Mathematical & Computational Sciences (SMCS)	**Kolkata, India**
Jul 2020 **Aug 2018**	**Indian Association for the Cultivation of Science** M.Sc. in Mathematics & Computing from SMCS	**Kolkata, India**
2018 **2015**	**Serampore College, University of Calcutta** B.Sc.(Hons.) in Mathematics	**Kolkata, India**

Research Contributions & Artifacts

DTECT: Dynamic Topic Explorer & Context Tracker 2025

- An **end-to-end interactive system** for exploring and interpreting the **evolution of topics** in **temporal corpora**, integrating **modeling**, **evaluation**, **visualization**, and **LLM-assisted insights**.
- **Artifacts:** GitHub Codebase | Hugging Face Demo | Video Walkthrough

S2WTM 2025

- A novel topic model, the **Spherical Sliced-Wasserstein Autoencoder**, that mitigates **posterior collapse** on **hyperspherical manifolds** to generate more coherent and diverse topics. Presented at **ACL 2025**.
- **Artifacts:** GitHub Codebase

GINopic 2024

- A topic modeling framework that leverages **Graph Isomorphism Networks (GIN)** to capture the **mutual dependencies between words**, an aspect often missed by contextualized models. Presented at **NAACL 2024**.
- **Artifacts:** GitHub Codebase

CTM-KD 2023

- A **knowledge distillation** framework to **compress** large topic models by minimizing both cross-entropy and the **2-Wasserstein distance** between latent distributions. Our student model surpasses the teacher in topic coherence with far fewer parameters. Presented at **ECIR 2023**.
- **Artifacts:** GitHub Codebase

CTM-Neg 2022

- An enhancement for contextualized topic models that introduces a **negative sampling** mechanism using a **triplet loss** on perturbed document-topic vectors. This approach significantly improves both **topic coherence** and **diversity**. Presented at **ICON 2022**.
- **Artifacts:** GitHub Codebase

Publications

J=Journal, C=Conference, W=Workshop, P=Preprint

[P] **DTECT: Dynamic Topic Explorer & Context Tracker** [🔗]
Suman Adhya and Debarshi Kumar Sanyal
Code | *Demo* | *Video* **[arXiv]**

[C] **S2WTM: Spherical Sliced-Wasserstein Autoencoder for Topic Modeling** [🔗]
Suman Adhya and Debarshi Kumar Sanyal
Association for Computational Linguistics, Vienna, Austria **[ACL'25]**

[J] **Evaluating Negative Sampling Approaches for Neural Topic Models** [🔗]
Suman Adhya, Avishek Lahiri, Debarshi Kumar Sanyal, and Partha Pratim Das
IEEE Transactions on Artificial Intelligence
An extension of work initiated under the guidance of Prof. Partha Pratim Das (Ashoka University). **[IEEE-TAI]**

[C] **GINopic: Topic Modeling with Graph Isomorphism Network** [🔗]
Suman Adhya and Debarshi Kumar Sanyal
North American Chapter of the Association for Computational Linguistics, Mexico City, Mexico **[NAACL'24]**

[C] **Do Neural Topic Models Really Need Dropout? Analysis of the Effect of Dropout in Topic Modeling** [🔗]
Suman Adhya, Avishek Lahiri, and Debarshi Kumar Sanyal
17th European Chapter of the Association for Computational Linguistics, Dubrovnik, Croatia **[EACL'23]**

[C] **Improving Neural Topic Models with Wasserstein Knowledge Distillation** [🔗]
Suman Adhya and Debarshi Kumar Sanyal
45th European Conference on Information Retrieval, Dublin, Ireland **[ECIR'23]**

[C] **Improving Contextualized Topic Models with Negative Sampling** [🔗]
Suman Adhya, Avishek Lahiri, Debarshi Kumar Sanyal, and Partha Pratim Das
19th International Conference on Natural Language Processing, New Delhi, India
This work was initiated under the foundational guidance of Prof. Partha Pratim Das (IIT Kharagpur). **[ICON'22]**

[W] **What Does the Indian Parliament Discuss? An Exploratory Analysis of the Question Hour in the Lok Sabha** [🔗]
Suman Adhya and Debarshi Kumar Sanyal
1st Workshop on Natural Language Processing for Political Sciences, LREC, Marseille, France **[PoliticalNLP @ LREC'22]**

Skills

Programming Languages	Python (Proficient), C, R (Familiar)
Tools & Frameworks	PyTorch, TensorFlow, HuggingFace, LangChain, Scikit-learn, Pandas, Scipy, Gensim, Sentence-Transformers, PyTorch Geometric (PyG), FAISS, Streamlit
Algorithms & Methods	Topic Modeling, Representation Learning, Knowledge Distillation, Negative Sampling, Optimal Transport, Graph Neural Networks (GNNs)

Presentations

- ACL 2025, Poster Presentation — July 2025 (Vienna, Austria)
- PIC 2025, Lighting Talk and Poster Presentation — Feb 2025 (Mysore, India)
- NAACL 2024, Poster Presentation — June 2024 (Virtual)
- PIC 2024, Poster Presentation — Feb 2024 (Mysore, India)
- EACL 2023, Poster Presentation — May 2023 (Dubrovnik, Croatia)
- ECIR 2023, Short Paper Presentation — March 2023 (Virtual)
- ICON 2022, Full Paper Presentation — December 2022 (IIIT-Delhi, India)
- PoliticalNLP @ LREC 2022, Poster Presentation — June 2022 (Virtual)

Travel Grants and Awards

ACM India-IARCS Travel Grant, 2025 [🌐] : Awarded for presenting our paper at ACL'25 held in Vienna, Austria.

ANRF-International Travel Support, 2025 [🌐] : Awarded for presenting our paper at ACL'25 held in Vienna, Austria.

ACM India Research Facilitation Grant, 2025 [🌐] : Awarded to cover the extra page charges of our paper at IEEE-TAI.

Microsoft Research Travel Grant, 2024 [🌐] : Awarded for presenting our paper at NAACL'24 held in Mexico City, Mexico.

ACM India-IARCS Travel Grant, 2024 [🌐] : Awarded for presenting our paper at NAACL'24 held in Mexico City, Mexico.

Travel Grant for Research Week with Google, 2024 [🌐] : Awarded for attending Research Week with Google, a three-day lineup of research lectures, panels, and discussions led by esteemed researchers.

Travel Grant for PIC, 2024 [🌐] : Awarded for attending flagship research networking event of ACM India held in the Infosys Mysore campus, India.

Travel Grant for CODS-COMAD, 2024 [🌐] : Awarded for attending the 7th Joint International Conference on Data Sciences and Management of Data to be held in IIIT Bangalore, India.

Microsoft Research Travel Grant, 2023 [🌐] : Awarded for presenting our paper at EACL'23 held in Dubrovnik, Croatia.

Travel Grant for EACL, 2023 [🌐] : Awarded for presenting our paper at EACL'23 held in Dubrovnik, Croatia.

The INSPIRE-SHE Award, DST, 2015-2018 [🌐] : Awarded for securing a top 1% rank in Class 12th examinations.

Teaching Assistant

Introduction to Machine Learning [PHD 226/ COM 4203], IACS January - June, 2024

- Assisted a graduate-level course with approximately **20 students**.
- **Responsibilities**: Designed and graded all programming assignments, and conducted dedicated doubt-clearing sessions to resolve student queries.
- **Topics Covered**: Assignments spanned key ML concepts, including Linear Regression, Decision Trees, SVM, k-NN, the EM Algorithm, and Neural Networks.

Laboratory - Numerical Methods (PHS 4211), IACS January - June, 2021

- Managed lab sessions for a cohort of **20+ master's students**.
- **Responsibilities**: Designed and graded weekly lab assignments in Python. Conducted tutorials and dedicated doubt-clearing sessions to support students with implementation.

Academic Responsibility

Volunteer coordinator	EACL'23
Review committee member	ARR - Feb'25, EMNLP'24 Industry Track, ARR - June'24, ARR - April'24, ARR - February'24, ACL'23, EMNLP'22
Program committee member	PoliticalNLP'24, EMNLP'23 Industry Track, PoliticalNLP'22

Community Service

Student Representative, Anti-Ragging Cell, IACS 2023 - Current
Actively contribute to fostering a safe and supportive campus environment for all students.

Research and Technical Assistant, Research Scholars' Association, IACS 2023-2024
Provided technical support and research guidance to fellow scholars within the association.

System Manager, Research Scholars' Association, IACS 2022-2023
Managed and maintained the IT infrastructure and online resources for the scholars' association.

References

References will be provided on request.