

# Suman Adhya

RA-I, School of Mathematical & Computational Sciences, IACS

[✉ adhyasuman30@gmail.com](mailto:adhyasuman30@gmail.com) | [🌐 adhyasuman.github.io](https://adhyasuman.github.io) | [👤 AdhyaSuman](https://www.linkedin.com/in/sumanadhy/) | [🔗 sumanadhy](https://www.linkedin.com/in/sumanadhy/)

## RESEARCH INTERESTS

---

Agentic AI • LLM Safety & Alignment • Topic Modeling • NLP • ML

## TECHNICAL SKILLS

---

Languages	Python (Proficient), C, R
GenAI Stack	RAG, LangChain, FAISS, Hugging Face, OpenAI/Gemini/Claude APIs
Frameworks	PyTorch, PyTorch Geometric (PyG), TensorFlow, Scikit-learn, Pandas
Tools & Viz	Streamlit, Git, Plotly, Sentence-Transformers, Gensim
Expertise	Generative Models, GNNs, Contrastive Learning, Knowledge Distillation

## EDUCATION

---

Oct, 2020 – Jan, 2026	<b>Ph.D., Computer Science, IACS</b> , India Thesis Title: <i>Studies on Neural Topic Models for Document Corpora</i>
Aug, 2018 - Jul, 2020	<b>M.Sc., Mathematics &amp; Computing, IACS</b> , India
2015 - 2018	<b>B.Sc. (Honours), Mathematics</b> , University of Calcutta, India

## KEY PROJECTS

---

### Generative AI Systems & Large-Scale Analysis

- **DTECT: Dynamic Topic Explorer & Context Tracker** [👤 (AAAI '26 Demo)]  
Built a unified end-to-end system for interpreting thematic evolution in text corpora. Developed a scoring method to identify temporally significant terms. Built a retrieval pipeline with FAISS and MMR for diverse document selection, supporting a grounded LLM chat assistant using LangChain.
- **Scalable Analysis of Parliamentary Data** [👤 (LREC '22)]  
Analyzed 298k Q&A pairs via dynamic topic models to quantify policy drift. Detected socio-economic shifts like 2007 credit card crisis and post-2016 digital payments transition (UPI).

### Representation Learning & Structured Knowledge

- **S2WTM: High-Fidelity Latent Representations** [👤 (ACL '25)]  
Leveraged Spherical Sliced-Wasserstein distance to model hyperspherical latent spaces, effectively mitigating posterior collapse. Outperformed Euclidean baselines with a 54.6% NPMI improvement and superior coherence validated by LLMs.
- **GINopic: Graph-Informed Text Modeling** [👤 (NAACL '24)]  
Fused TF-IDF representations with graph-structured knowledge using Graph Isomorphism Networks (GIN). This dual-stream architecture captures global semantic correlations, achieving state-of-the-art coherence and classification accuracy across 5 benchmark datasets.
- **Contrastive Learning Framework for Topic Models** [👤 (IEEE TAI '25)]  
Designed a model-agnostic Decoder Negative Sampling (VAE-DNS) framework compatible with 7 neural architectures. Implemented Triplet Loss to enforce latent space disentanglement, achieving superior cluster separation on Wiki40B (verified via UMAP).

## Model Compression & Efficiency

- **Wasserstein Knowledge Distillation for Topic Models** [Q] (ECIR '23)  
Engineered a compression framework to distill a hybrid Teacher (BoW + SBERT) into a streamlined SBERT-only Student. Utilized Squared 2-Wasserstein distance to align latent geometries, achieving a model size reduction of 37.6% – 56.3% (55.4% on 20NG) while enabling the compressed student to surpass the teacher's topic coherence.
- **Regularization Dynamics in Generative Models** [Q] (EACL '23)  
Demonstrated that high dropout disrupts manifold learning in generative models like VAEs. Established low-dropout protocols ( $p < 0.2$ ) as essential for preserving generation quality in VAEs.

## SELECTED PUBLICATIONS

---

For a complete list, see my [Google Scholar profile](#).

- **Suman Adhya** and Debarshi Kumar Sanyal. [DTECT: Dynamic Topic Explorer & Context Tracker](#). **AAAI**, 2026 (Demo). [Q] · [G]
- **Suman Adhya** and Debarshi Kumar Sanyal. [S2WTM: Spherical Sliced-Wasserstein Autoencoder for Topic Modeling](#). **ACL**, 2025. [Q]
- **Suman Adhya**, Avishek Lahiri, Debarshi Kumar Sanyal, and Partha Pratim Das. [Evaluating Negative Sampling Approaches for Neural Topic Models](#). **IEEE TAI**, 2025. [Q]
- **Suman Adhya** and Debarshi Kumar Sanyal. [GINopic: Topic Modeling with Graph Isomorphism Network](#). **NAACL**, 2024. [Q]
- **Suman Adhya**, Avishek Lahiri, and Debarshi Kumar Sanyal. [Do Neural Topic Models Really Need Dropout? Analysis of Dropout in Topic Modeling](#). **EACL**, 2023. [Q]
- **Suman Adhya** and Debarshi Kumar Sanyal. [Improving Neural Topic Models with Wasserstein Knowledge Distillation](#). **ECIR**, 2023. [Q]

## GRANTS AND SCHOLARSHIPS

---

- **Travel Grants (ACL 2025)**: ACM India-IARCS & ANRF-ITS. 2025
- **ACM India Research Facilitation Grant**: Support for IEEE TAI publication. 2025
- **Travel Grants (NAACL 2024)**: MSR Grant & ACM India-IARCS 2024
- **Travel Grants (EACL 2023)**: MSR Grant, D&I Team, & Volunteer Coordinator. 2023
- **INSPIRE-SHE Award (DST)**: For top 1% rank in National Class 12th Exams. 2015–2018

## COMMUNITY SERVICE

---

- Student Representative, Anti-Ragging Cell, IACS 2023 - Current
- Research and Technical Assistant, Research Scholars Association, IACS 2023 - 2024
- System Manager, Research Scholars Association, IACS 2022 - 2023