Suman Adhya

AI/NLP Researcher

Jadavpur, Kolkata-700032, West Bengal, INDIA

+91 9007657578 | adhyasuman30@gmail.com | AdhyaSuman | sumanadhya

Skills & Expertise

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

Professional Experience

Research Associate-I | *SMCS, IACS* Jan 2026 – Present

– **LLM Safety & Alignment:** Collaborating on research to develop safety guardrails and alignment methods for LLMs.

– **Agentic AI Evaluation:** Designing frameworks for trajectory-level evaluation to evaluate the planning, reasoning, and multi-step execution of AI agents.

AI Researcher (Ph.D. Candidate) | *SMCS, IACS* Oct 2020 – Jan 2026

Publication Venues: *AAAI, ACL, NAACL, IEEE TAI, EACL, ECIR, LREC (workshop).*

– Spearheaded end-to-end applied research in Generative AI, representation learning, and system architecture (detailed in Key Projects below).

Education

Oct, 2020 – Jan, 2026	**Ph.D., Computer Science**, IACS, India Thesis Title: *Studies on Neural Topic Models for Document Corpora*	Defended: *15/01/2026*
Aug, 2018 – Jul, 2020	**M.Sc., Mathematics & Computing**, IACS, India	
2015 – 2018	**B.Sc. (Honours), Mathematics**, 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 the 2007 credit card crisis and the post-2016 digital payments transition (UPI).

Model Compression & Efficiency

– **Wasserstein Knowledge Distillation for Topic Models** [(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 55.4% on 20NG while enabling the compressed student to surpass the teacher's topic coherence.

– **Regularization Dynamics in Generative Models** [(EACL '23)]
Empirically demonstrated that high dropout rate disrupts representation learning in VAEs, establishing low-dropout protocols ($p < 0.2$) as essential for preserving generation quality in VAEs.

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).

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

Awards & Leadership

- **Competitive Research & Travel Grants:** Awarded multiple grants from ACM India, ANRF, and Microsoft Research to present at top-tier conferences (AAAI '26, ACL '25, EACL '23).
- **INSPIRE-SHE Award (DST):** Nationally recognized for ranking in the top 1% in Class 12th Board Exams.
- **Institutional Leadership (IACS):** Served as Student Representative for the Anti-Ragging Cell, and as Research and Technical Assistant and System Manager for the Research Scholars Association.