

Trishan Singha Roy

📍 Delhi, India ✉ Email 🌐 Website in LinkedIn 🐙 GitHub

Education

Indian Institute of Technology, Delhi

2023 - Present

B.Tech in Electrical Engineering | CGPA : 9.03/10

Department Rank : 18

Delhi Public School, Ruby Park

2021 - 2023

CBSE Class XII | 94.8%

Internships / Research Experience

IBM, Software Innovation Lab | Storage-Efficient Multimodal Document Retrieval

May 2026 - July 2026

- Developed **ColSNAP**, a training method for **multimodal RAG** that reduces the storage cost of visual document retrievers while enabling flexible accuracy-storage trade-offs at indexing time (under review at **AAAI 2027**).
- Reduced vector-embedding storage cost by **86.3% (7.3×**) with lossless retrieval performance, and up to **98.6% (69.2×**) while retaining **98.3%** of full-resolution, uncompressed performance.
- Introduced a novel **Spatial Matryoshka Loss** to jointly train a single model across multiple compression levels with minimal additional training; demonstrated that the approach transfers effectively across architectures.

LCS2 Lab, IIT Delhi | Improving Temporal Consistency in Agentic LLMs

Oct 2025 - Present

Advisor : Prof. Tanmoy Chakraborty

- Built a benchmark for **temporal reasoning consistency** in tool-using LLMs, evaluating factuality and answer invariance across semantically equivalent paraphrases of date, event, and temporal-relation queries.
- Developing a novel **paraphrase-aware RL objective** that transfers reward signals across equivalent temporal queries to improve agentic reasoning consistency, outperforming **GRPO** and **PPO** baselines.

CogAI4Sci, NUS Singapore | Latent Reasoning with LLMs

May 2025 - Aug 2025

Advisor : Prof. Dianbo Liu

- Explored **latent reasoning** through discrete and continuous latent token representations, experimenting with RL-based training approaches including **GRPO** for latent token learning, matching baseline CoT accuracy on **GSM8K**.
- Investigated and experimented with various **latent token initialization** techniques, including training and analysing a sentence-encoder based **CoT compression framework** to support model interpretability and efficiency.
- Quantified the novelty and LLM usage of published research and studied the correlation between them, utilizing keyword and reference embeddings.

LCS2 Lab, IIT Delhi | Multi-Agent LLM Collaboration for Long-Horizon Tasks

Aug 2026 - Present

Advisor : Prof. Tanmoy Chakraborty

- Analyzing the impact of worker-agent time constraints on overall **multi-agent system** performance in long-horizon tasks.
- Developing a novel **error-injection framework** that deliberately corrupts agent context to evaluate system-level reliability, error recovery, and robustness over long-horizon tasks.

Publications

ColSNAP: Spatial Matryoshka Training for Multi-Granularity Visual Document Retrieval [Under Review, AAAI 27]

Trishan Singha Roy, Arkadeep Acharya, Vishwajeet Kumar, Jaydeep Sen, Sachindra Joshi [arXiv]

- Introduced a novel multi-granularity training framework for visual document retrieval that learns a hierarchy of nested compressed representations, enabling flexible storage-accuracy trade-offs at deployment.

Projects

Safe and Data-Efficient Reinforcement Learning

GitHub

Prof. Raunak Bhattacharya

- Engineered a Safe Reinforcement Learning pipeline based on constrained PPO, implementing primal-dual optimization, cost-aware advantage estimation, adaptive constraint penalties, and stability mechanisms entirely from scratch.
- Characterized safety–reward trade-offs in constrained continuous control, conducting multi-seed 1M-step experiments to quantify reward–safety trade-offs and identify robust constraint-enforcement strategies.
- Investigated data efficiency in Offline RL, constructing decoupled value/policy training pipelines and evaluating **IQL**, **TD3+BC**, and **CQL+AWR** across systematic critic–actor dataset scaling regimes.

GPT2 based Context Aware Decoder

GitHub

- Enhanced GPT-2's generation capabilities by integrating context-aware decoding (CAD) to improve faithfulness and **reduce hallucination** in generated content by the method described in Shi et al.
- Modified the forward pass of the GPT-2 model to incorporate CAD, utilizing a contrastive output distribution to amplify **context-relevant probabilities**.
- Addressed issues of context-prior conflicts in text generation by implementing a decoding strategy that emphasizes external contextual information.

CharacterCNN based Hangman Solver

GitHub

- Developed a lightweight CharacterCNN model leveraging character-level convolutions to capture subword patterns and predict masked letters in Hangman.
- Designed a synthetic dataset generator simulating masked words and incorrect guesses for supervised training.
- Implemented a GLU-based convolutional stack with attention pooling and contextual encoding of wrong guesses, achieving an improvement of **26%** over BERT based methods.

SRTF Algorithm for Dynamic Resource Allocation and Processing

GitHub

Prof. Amit Kumar

- Designed and implemented a scheduling system based on the Shortest Remaining Time First (SRTF) algorithm to manage dynamic task distribution in a parallel computation environment.
- Developed a custom min-heap data structure to efficiently prioritize and assign tasks according to user-defined criteria, ensuring optimal resource allocation.
- Built a scalable solution to handle multiple calls and dynamic task updates, ensuring real-time adjustments and minimal processing delays.

Achievements

- Citadel India Terminal** : Finished **6th**, out of hundreds of entrants all over India.
 - Created an algorithm to play a custom turn-based Tower Defence game, mathematically optimizing the balance of attack and defence under cost constraints while accounting for algorithm speed.
- IBM Best Poster Award** : Awarded **Best Poster** by a panel of judges, out of **70+** interns.
- Department Rank** : **18** out of **200+** students in the Electrical Department.
- Regional Mathematics Olympiad** : Qualified for the **RMO**, ranking among the top **1%** of candidates statewide.
- WBJEE 2023** : Secured **Rank 80** amongst **120,000+** candidates.
- Department Change** : Among top **50** out of **1200+** students to secure department change on CGPA basis after 1st year.

Technical Skills

- Languages:** Python, C++, SQL
- Tools and Libraries:** NumPy, Pandas, Scikit-learn, PyTorch, TensorFlow, OpenCV, Hugging Face, vLLM, Docker
- Certifications:** Coursera's Deep Learning Specialisation including - Convolutional Neural Networks, Sequence Models, Hyperparameter Tuning, Regularization and Optimization