

RISHABH SABHARWAL

+44 7486 045605 | R.Sabharwal@sms.ed.ac.uk

[Website](#) | [LinkedIn](#) | [Github](#) | [Scholar](#)

OBJECTIVE

I am a master's student at the University of Edinburgh, supervised by [Prof. Jeff Pan](#). For my thesis, I am creating a benchmark to assess **Self-evolving LLM Agents** that integrate web search and coding into end-to-end workflows. Previously, I was a Predoctoral Researcher at the Indian Institute of Science (IISc), working under [Prof. Punit Rathore](#) on projects related to graph neural networks and computer vision.

EDUCATION

MSc in Artificial Intelligence, University of Edinburgh, UK (2025–2026)

B.Tech in ECE (AI & ML), Netaji Subhas University of Technology, India (2020–2024) | **GPA: 3.68/4**

PUBLICATIONS

A.Mazumder, **Rishabh Sabharwal**, M.Tayal, B. Kumar, C. Garg, P. Rathore, ‘On Convergence of Adam with Data Dependent Stepsize’. (Accepted at **IEEE Transactions on AI**, to be published.)

Rishabh Sabharwal, P.S. Rathore, P. Rathore, “DiGCT: Diffusion Guided Gaussian Context Transformer”.

R. Samarth*, **Rishabh Sabharwal***, P. Rathore, “SSGNN: Simple Yet Effective Spectral GNN”. ([OpenReview](#))

Rishabh Sabharwal, S. Syal, P. Pankaj, “Optimizing Stock Market Predictions: A 3-Stage Model with Residual Modelling”. In Proceedings of IEEE Signal Processing and Integrated Networks, 2024. ([Paper](#))

SELECTED PROJECTS

FastAPI Safety Gateway: ([Code](#))

- Built a guardrail around a vLLM-served Llama-3.1-8B-Instruct, combining learned and rule-based guardrails like multi-label toxicity/harm classification, prompt-injection detection, and PII detection/anonymization.
- Developed an **agentic safety loop** that rewrites risky prompts and repairs unsafe outputs through policy conditioned re-prompting.
- Enhanced deployability through user-adjustable thresholds, custom deny lists and transparent decision tracing via the Streamlit UI.

Video Humor Reasoning:

- Developed the **Indic-SMILE** dataset to generate plausible reasons behind humor in video clips. Classified clips into different humor categories and extracted audio, visual and textual features from the clips.
- Benchmarked VLMs and LLMs (InternVL3, VideoLLaMa3 Gemini-2.5 Flash, Claude 4.5 Sonnet, Gemma-3, Sarvam-M) to evaluate **cultural alignment** in Indian contexts.

Face X-ray for Forgery Detection: ([Code](#))

- Developed a model that can locate the blending boundary (face-x-ray) in forged images. Created a set of blended images and corresponding boundaries from the FFHQ dataset.
- DeepLabV3 semantic segmentation model with EfficientNet as the encoder was finetuned to achieve 94% accuracy on the test set. A similar project to **detect deepfake images** was completed with Omdena. ([Presentation](#))

EXPERIENCE

Pre-doctoral Researcher

Indian Institute of Science

Jun 2024 - Jul 2025

Bengaluru, India

- Worked with **NPCI** (National Payments Corporation of India) to build efficient and scalable large-scale GNNs to detect fraudulent transactions on graphs having more than a **billion nodes**.
- Developed an efficient spectral GNN featuring a novel, parameter-free module called ReGA (Relative Gaussian Amplifier) that performs global context filtering. **Evaluated on 20 real-world graph datasets**; achieved SOTA performance, reducing parameters by an average of 55x and GFLOPs by 100x across all datasets.

- Extended the Gaussian Context Transformer (GCT) to develop DiGCT, utilizing Graph Tikhonov Regularization and graph diffusion for stronger channel attention. Increased average accuracy by **2%** on modern ConvNets with **only 20 additional parameters**.

Research Intern

Indian Institute of Science

Jan 2024 - Jun 2024

Bengaluru, India

- Proposed a network and data-dependent constant learning rate (lr) to address the issues of decaying lr schedulers and reduce the need for extensive lr search.
- Conducted empirical analysis of our proposed lr which results in faster convergence for both stochastic and deterministic variants of the Adam optimizer. **Paper accepted at IEEE Transactions on AI.**

Research Intern

Indraprastha Institute of Information Technology

May 2023 - Jul 2023

Delhi, India

- Developed a speech-to-text model to streamline the patient registration process in government hospitals, tackling the problems of **different dialects and frequent language switching** between Hindi and English.
- Fine-tuned and benchmarked Whisper, Meta MMS, VOSK, NVIDIA Nemo models on Indian-accented English and Hindi datasets and some internal medical-speech samples.
- Developed a pipeline that includes a voice activity detector to reduce hallucinations in long audio samples, along with a quantized and fine-tuned Whisper model.
- Our prototype secured **Top-10 rank in the nation-wide Bhashini Grand Innovation Challenge.**

RELEVANT COURSEWORK

- **Accelerated NLP:** Classical and Neural LMs; Attention/Transformers (BERT/T5/GPT); Chain-of-Thought (CoT) Reasoning; Scaling laws; Post-training (instruction tuning, RLHF).
- **Probabilistic Modelling and Reasoning:** Graphical Models; Exact inference (DAGs/HMMs); Causality and Decision-making under uncertainty; Latent variable models; Variational inference and VAEs; Monte Carlo sampling.
- **Programming for Data Science at Scale:** Data-parallel programming model; distributed data-parallel systems; sparse processing; optimization of distributed pipelines; distributed query and graph processing.
- **Current Subjects:** Advanced Topics in ML, Machine Learning Systems and Methods in Causal Inference.

SCHOLASTIC ACHIEVEMENTS

- Recipient of **Kotak-IISc pre-doctoral fellowship** given by Kotak-IISc AIML centre. ([website](#))
- Selected as **Amazon MLSS** mentee, a 4 week program conducted by Amazon's research scientists.

ONLINE CERTIFICATIONS

- **Oxford ML Summer School:** Representation Learning and Generative AI (July 2024)
- **DeepLearning.AI:** Deep Learning Specialization (Feb 2023), [Machine Learning Specialization](#) (July 2022)
- **IBM:** Data Science Professional Certificate (Dec 2022) ([Code files](#))
- **Udemy:** Data Structures and Algorithms in Python (July 2022)

SKILLS

- **Programming Languages:** Python, C++, Scala, SQL
- **Frameworks:** PyTorch, PyG, JAX, Keras, FastAPI, Streamlit
- **ML/LLM Tooling:** Hugging Face, vLLM, OpenAI-compatible APIs
- **DevOps/Systems:** Git, Linux, CI/CD, PyTorch DDP (NCCL)