

Priyanshu Raj Mall

Email: mall.priyanshu07@gmail.com

Phone: +91-80949-17702

GitHub: github.com/mallpriyanshu

LinkedIn: [Priyanshu Raj Mall](#)

EDUCATION

- **Indian Institute of Science Education and Research (IISER) Kolkata** India
BS-MS Major: Physics, Minor: Computational and Data Sciences; CGPA: 8.43/10 2018 – 2023
Relevant Courses: Information Retrieval and Web Search, Natural Language Processing, Machine Learning and Network Analysis, Computational Physics, Research Methodology, Programming and Data Structures I/II, Probability I, Linear Algebra I, Mathematical Methods I/II, Introduction to Computer Programming

SKILLS SUMMARY

- **Languages** Python, C (familiar), Java (familiar), MATLAB, HTML/CSS, SQL
- **Frameworks and Libraries** Transformers, PyTorch, Pyserini, Gensim, TensorFlow, PyLucene, NumPy, SciPy, Faiss, Matplotlib, Pandas, Scikit-learn, django
- **Tools & Platforms** Linux, AWS, Pinecone, Git, GitHub, LaTeX, Origin, GnuPlot, ImageJ
- **Analytics** Hypothesis Testing, Statistical Analysis (Regression, t-test, Chi-square)

EXPERIENCE

- **SIGIR 2025 LiveRAG Challenge** Remote, Italy
RAG Pipeline Development Mar 2025 – Jul 2025
 - **Hybrid RAG System:** Built modular RAG pipeline combining dense (Pinecone) and sparse (OpenSearch) retrieval using Reciprocal Rank Fusion (RRF).
 - **LLM Integration:** Integrated Falcon and other LLMs via AWS Bedrock for chunk extraction and answer generation.
- **Datathon@IndoML 2024 (Sponsored by NielsenIQ)** Goa, India
Winner – National ML Competition Aug 2024 – Dec 2024
 - **Ensemble Modeling:** Developed multi-model ensemble combining ByT5, LoRA-LLaMA + Flan-T5, multi-modal ByT for attribute-value prediction on noisy e-commerce data.
- **Postgraduate Research Assistant – IISER Kolkata** Kolkata, India
Research Student Under Dr. Dwaipayan Roy Jul 2023 – May 2024
 - **Research Output:** Contributed to 3 publications in top-tier IR venues (CIKM'23, FIRE'23, ECIR'24).
 - **Core Contributions:** Designed retrieval experiments, analyzed retrievability bias, and built evaluation pipelines.
- **Teaching Assistant @ IISER Kolkata** Kolkata, India
Undergraduate and Graduate Courses 2021 – 2023
 - Assisted in **Information Retrieval and Web Search (CS4201)**, **Introduction to Computer Programming (CS1101)**, and **Physics Laboratory II (PH1202)** over five semesters.
 - Supported lab sessions, evaluated assignments, and guided students in Python programming and experiments.

PUBLICATIONS AND TECHNICAL REPORTS

[[GOOGLE SCHOLAR](#)]

- **SIGIR'25** – Hybrid Retrieval-Augmented Generation Pipeline for Diverse QA Tasks [Link](#)
- **INDOML'24** – Attribute-Value Prediction for E-Commerce: A Multi-Model Ensemble Approach [Link](#)
- **ECIR'24** – Exploring the Nexus Between Retrievability and Query Generation Strategies [paper](#)
- **FIRE'23** – A Comparative Analysis of Retrievability and PageRank Measures [paper](#)
- **CIKM'23** – Findability: A Novel Measure of Information Accessibility [paper](#)

PROJECTS

- **RAG Pipeline for LiveRAG Challenge (SIGIR '25)** Developed a dense+sparse hybrid RAG pipeline using Pinecone, OpenSearch, RRF fusion, and neural reranking. Integrated LLMs for modular, cloud-based QA system.
- **Pattern Classification using Extra Trees** Achieved 11.5% improvement over baseline on medical image datasets using Extra Trees Classifier and Gini feature network analysis.
- **Hopfield Neural Networks for Associative Memory** Implemented Ising-model-inspired Hopfield net for content-addressable memory and partial-image reconstruction.
- **Rumor Spread Modeling** Simulated and analyzed rumor dynamics using a modified SIR-based model with control parameters, comparing spread behavior under interventions.

HONORS AND AWARDS

- **Winner, Datathon@IndoML 2024** – National-level ML competition sponsored by NielsenIQ; presented winning solution at IndoML 2024 Symposium.
- **LiveRAG Challenge, SIGIR 2025** – Awarded \$1,500 AWS and \$700 Pinecone credits; selected for oral presentation at ACM SIGIR LiveRAG 2025 workshop in Padua, Italy.
- **Admitted Participant, Lisbon Machine Learning School (LxMLS) 2025** – Selected from 450+ applicants based on research and ML background.

CERTIFICATIONS

- **Scripting in Python Specialization** – Rice University Completed via Coursera; focused on automation and Python scripting. [certificate](#)
- **The Data Scientist's Toolbox** – Johns Hopkins Introductory course on data science tools and workflows. [certificate](#)