Mandeep Rathee

Hannover, Germany | mandeeprathee77@gmail.com | mandeep-rathee.github.io

linkedin.com/in/mandeep-rathee | github.com/Mandeep-Rathee

About myself

I am experienced in **machine learning** and **deep learning** techniques, applied in diverse domains such as **search, ranking, natural language processing, and graphs.** I have worked with search and ranking tools such as **Pyserini and Pyterrier.** I am passionate about applying my research findings to address real-world challenges. This involves developing advanced and reliable models that can effectively tackle practical issues and make a positive impact in society.

Technologies

Python, PyTorch, Hugggingface, Numpy, Git, Bash, Pandas, Matplotlib, Scikit-learn, Flask-restx, R, Docker, Swagger, SQL, Pyserini, Pyterrier, Langchain

Education

PhD Researcher, L3S Research Center, Hannover, Germany — Sept 2020 – Present

- Working on solving the bounded recall problem in standard retrieve and rank (also termed as cascading) pipelines.
- Proposed novel relevance estimation methods for adaptive **retrieval** and **ranking**.

M.Tech., Indian Institute of Technology, Patna, India — July 2018 – Aug 2020

- Worked on improving the effectiveness of the graph machine learning models.
- Got DAAD fellowship for master's thesis at L3S Research Center, Hannover.

Experience

Research Assistant, L3S Research Center, Hannover — Jan 2022 - Present

- Developed and implemented algorithms end-to-end to improve **search** and **recommendation** results for a digital learning platform MLS under the project called SEARCH, fully funded by the **BMBF, Germany.**
- Developed an API called AIMS for **MLS** which helps content creators in generating metadata using **Large Language Models.**
- Worked closely with the front-end engineering team to integrate new APIs and **search** & **recommendation** functions into the learning platform.

Research Project

Breaking the Lens of the Telescope: Online Relevance Estimation over Large Retrieval Sets, in SIGIR, 2025

- Proposed an **Online Relevance Estimation (ORE)** framework that learns the scores of the expansive reranker for unranked documents.
- The ORE framework can be used with hybrid and adaptive retrieval settings without adding latency overheads.

Paper: `https://arxiv.org/pdf/2504.09353`

SUNAR: Semantic Uncertainty based Neighborhood Aware Retrieval for Complex QA, in NAACL, 2025

- Propsed noval RAG pipeline using adaptive retrieval based on the corpus graph for complex QA.
- Our proposed method outperforms existing state-of-the-art methods (including Self-RAG and Self-Ask) by up to 31% in performance.

Paper: `https://arxiv.org/pdf/2503.17990` Code: `https://github.com/VenkteshV/SUNAR`

Guiding Retrieval using LLM-based Listwise Rankers, in ECIR, 2025

- Proposed Slide window-based adaptive retrieval method called SlideGAR to improve retrieval using LLM-based listwise rankers.
- Proposed method improves nDCG@10 by 13% and Recall by 28% over state-of-the-art listwise rankers without adding latency overheads.

Paper: `https://arxiv.org/pdf/2501.09186` Code: `https://github.com/Mandeep-Rathee/llmgar`

Quam: Adaptive Retrieval through Query Affinity Modelling, in WSDM, 2025

- Proposed an **adaptive retrieval** approach based on the corpus graph, which improves recall by up to 26% for standard re-ranking systems.
- The Proposed algorithm is both efficient and effective over other baselines.

Paper: `https://doi.org/10.1145/3701551.3703584` Code: `https://github.com/Mandeep-Rathee/quam`