

GAURAV GUPTA

Applied Scientist II
Amazon Search,
Palo Alto, CA, United States

Webpage- <https://gaurav16gupta.github.io/>
Email- gauravgupta16may93@gmail.com
Github- <https://github.com/gaurav16gupta>

EDUCATION

Rice University, Houston TX, United States

(Aug 2018 - Aug 2023)

Ph.D. in Large scale Information Retrieval ([Thesis](#))

Advised by Prof. Anshumali Shrivastava

Indian Institute of Technology, Hyderabad, India

(Aug 2010 - May 2014)

Bachelor of Technology, Major in Electrical Engineering

INTERESTS AND SKILLS

Recommendation Models, Search, Ranking,, NLP, Large Language Models, Retrieval Augmented Generation, Computer Vision, Machine Learning, Generative AI, Deep Learning, HPC: OpenMP, CUDA

Languages/Libraries: Python, C++, PyTorch, TensorFlow

PROFESSIONAL EXPERIENCE

Amazon Search, Palo Alto, Applied Scientist II

(Nov 2023 - Present)

Manager – Tanya Roosta (Sr. Applied Scientist)

Product retrieval for Amazon search (Shopping)

1. Designed and implemented models for personalized search by leveraging customer behavior data. Optimized embedding models and ANN indices to deliver relevant and tailored search results for shopping.
2. Optimized and A/B tested an embedding-based search index. This improved latency and ensured seamless integration with existing search systems, enhancing user engagement and sales performance.
3. A/B tested query expansion/reformulations, resulting in significant revenue growth within a key market.
4. Presented research talk at an internal conference.
5. Collaborated with multiple teams and engineers to deliver high-impact solutions.

Adobe Research, San Jose, Research Intern

(June 2022 - Aug2022)

Manager – Dr. Vishy Swaminathan (Sr. Principal Scientist)

Worked on data drift detection for recommendation models.

Adobe Research, San Jose, Research Intern

(June 2021 - Aug2021)

Manager – Dr. Vishy Swaminathan (Sr. Principal Scientist)

Proposed a real-time product recommendation framework with dense-semantic and sparse-lexical features while beating the state-of-the-art methods on retrieval time. (Work published in Big Data, 2023)

Amazon Web Services (AWS), Palo Alto, Applied Scientist Intern

(June 2020 - Aug2020)

Manager – Dr. Alexander J. Smola (VP/Distinguished Scientist)

Proposed a Billion scale near neighbour index with a balanced load. Reduced the number of distance computations while beating state-of-the-art methods like HNSW. (Work published in SIGKDD 2022)

Amazon Search Labs (A9), Palo Alto, Applied Scientist Intern

(May 2019 - Aug2019)

Manager - Vijai Mohan (Senior Principal Applied Scientist)

Worked on efficient sublinear and low-memory indexing methods for constraint filtering. Proposed a novel idea of repeated and merged Bloom filter.

TCS Innovation Labs, Delhi, Full-time Researcher R&D

(July 2014 - Aug 2018)

Manager - Dr. Gautam Shroff (VP, TCS Fellow), Mentor - Dr. Lovekesh Vig

1. Worked on document reading, object recognition and classification for retail businesses, marker-based head motion tracker, indoor localization algorithms, and reconfigurable Inspection framework for augmented reality devices.
2. Mentored and worked with 2 interns and onboarded new hires in the team.
3. Published 5 papers in top-tier conferences.

CONFERENCE PUBLICATIONS

1. IDentity with Locality: An ideal hash for efficient gene sequence search, KDD 2025
*Gaurav Gupta**, *Tianyi Zhang**, *Aditya Desai**, *Anshumali Shrivastava*
2. CAPS: A Practical Partition Index for Filtered Similarity Search, ISIR-eCom, WSDM 2024
Gaurav Gupta, *Jonah Yi*, *Benjamin Coleman*, *Chen Luo*, *Vihan Lakshman*, *Anshumali Shrivastava*
3. BLISS: A Billion scale Index using Iterative Re-partitioning, KDD 2022
Gaurav Gupta, *Tharun Medini*, *Anshumali Shrivastava*, and *Alexander J. Smola*
Invited talk at NeurIPS 2021 BigANN competition track <https://big-ann-benchmarks.com>
4. Fast Processing and Querying of 170TB of Genomics Data via a Repeated And Merged BloOm Filter (RAMBO), 2226-2234, ACM SIGMOD, 2021
*Gaurav Gupta**, *M. Yan**, *B. Coleman*, *R. A. Elworth*, *T. Medini*, *T. Treangen*, *A. Shrivastava*
Featured News Article: RAMBO speeds searches on huge DNA databases,
<https://www.sciencedaily.com/releases/2021/06/210628152920.htm>
5. To Petabytes and beyond: recent advances in probabilistic and signal processing algorithms and their application to metagenomics, Nucleic acids research 48 (10), 5217-5234
R A Leo Elworth, *Qi Wang*, *Pavan K Kota*, *C J Barberan*, *Benjamin Coleman*, *Advait Balaji*, *Gaurav Gupta*, *Richard G Baraniuk*, *Anshumali Shrivastava*, *Todd J Treangen*
6. Near Neighbour Search for Constraint Queries , IEEE International Conference on Big Data, 2023,
Gaurav Gupta, *Anup Rao*, *Tung Mai*, *Ryan A Rossi*, *Xiang Chen*, *Saayan Mitra*, *Anshumali Shrivastava*
7. Siamese Networks for chromosomes classification, BIC, ICCV 2017 (pp. 72-81)
Swati, *Gaurav Gupta*, *M. Yadav*, *M. Sharma*, *L. Vig*
8. Information Extraction from Hand-marked Industrial Inspection Sheets, ICDAR 2017 (vol 6 pp 33-38) IEEE
Gaurav Gupta, *Swati*, *M. Sharma*, *L. Vig*
9. Indoor Localisation and Navigation on Augmented Reality Devices, IEEE ISMAR, 2016
Gaurav Gupta, *N. Kejriwal*, *P. Pallav*, *E. Hassan*, *S. Kumar*, *R. Hebbalaguppe*
10. An AR Inspection Framework: Feasibility Study with Multiple AR Devices, IEEE ISMAR, 2016
*Gaurav Gupta**, *R. Perla**, *R. Hebbalaguppe**, *G. Sharma*, *E. Hassan*, *M. Sharma*, *L. Vig*, *G. Shroff*
11. Robust Hand Gestural Interaction for Smartphone Based AR/VR Applications, 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), 2017, pp. 330-335, doi: 10.1109/WACV.2017.43
S. Mohatta, *R. Perla*, *Gaurav Gupta*, *E. Hassan* and *R. Hebbalaguppe*,

* equal authorship

PATENTS

1. System and methods for detecting and tracking a marker, US Patent 20,170,278,266
Gaurav Gupta, *Ehtesham Hassan*
2. Method and system for extracting information from hand-marked industrial inspection sheets, EP Patent 201,721,039,681, *Gaurav Gupta*, *Swati Jindal*, *Monika Sharma*, *Lovekesh Vig*
3. Sparse embedding index for search, US Patent US20230153338A1
Gaurav Gupta, *Anup Rao*, *Tung Mai*, *Ryan A. Rossi*, *Xiang Chen*, *Saayan Mitra*

INVITED TALKS/ PRESENTATIONS

In addition to conference paper presentations:

1. **Tutorial:** How to Deal with Volume and Velocity Associated with Hundreds of Terabytes (and Beyond) of Genomics Data? - 2021 Ken Kennedy AI and Data Science Conference.
2. **Presentation:** Repeated And Merged BloOm Filter: Querying on 170TB of Genomics Data at
 - Genome Informatics, 2020
 - Rice Oil & Gas High Performance Computing conference, 2020

RELEVANT COURSEWORK

Ph.D. (Rice University, Houston TX): Probabilistic Data structures, Multi-Core Processing, Machine Learning, NLP, Optimization, Networks and Graph theory, Random Processes, Statistical Signal Processing, Computational Imaging

POSITION OF RESPONSIBILITIES – RICE UNIVERSITY

1. Mentor for 2 undergraduate summer interns at RUSH lab, Rice University (Summer 2022)
2. Teaching assistant- Introduction to Deep Learning (Fall 2020)
(Prof. Ankit Patel, ELEC-COMP 516, ~80 students)
3. Teaching assistant- Probabilistic Algorithms and Data-structures (Spring 2020)
(Prof. Anshumali Shrivastava, COMP 480/580, ~50 students)
4. Reviewer at AISTAT-2021, CIKM-2022, NeurIPS-2022, ICLR 2022, ICML 2023, KDD 2023, NeurIPS 2023
5. Student volunteer at NeurIPS-2019, Vancouver, Canada (2019)
6. Academic Mentor for Rice ECE Graduate Student Association (2020)
7. Media Head (photography and videography) at Indian Students at Rice (ISAR) (2019-20)

POSITION OF RESPONSIBILITIES – IIT HYDERABAD

1. Student Coordinator at Training and Placement Cell. (2013-14)
2. Coordinator of Electronics Club. (2012-13)
3. Represented IIT Hyderabad at national level ABU ROBOCON- 2013 Pune. (2013)
4. Volunteer at National Service Scheme Team-IIT Hyderabad (2011)

OTHER ACHIEVEMENTS

1. Travel Grants –
 - MLSys 2022 (NSF Funded),
 - NeurIPS 2019.
2. Ph.D. panellist at undergrad student mentorship program at KDD 2022.
3. Nominated by the ECE department at Rice University for Google Ph.D. fellowship.
4. Ranked first at Neilsen's Image Recognition Hackathon- 2017 for "Hole detection/ product out of stock" retail store challenge.
5. Ranked 1913 (among top 0.4 %) in IIT-JEE engineering entrance exam, 2010.
6. Selected for Deep Learning Summer School- MILA, Montreal, 2016.