GAURAV GUPTA
Electrical and Computer Engineering
Rice University, Houston, TX, United States
Webpage- gaurav16gupta.github.io
Email- gg29@rice.edu

EDUCATION

Rice University, Houston TX, United States **(Aug 2018 - Expected 2023)**
Ph.D. in Information Retrieval and Large-Scale Machine Learning
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

Hashing, Information Retrieval, Learning to Index, Natural Language Processing, Large Language Models, Retrieval Augmented Generation, Computer Vision, Machine Learning, Deep Learning, HPC: OpenMP, CUDA
Languages/Libraries: Python, C++, PyTorch, TensorFlow

PROFESSIONAL EXPERIENCE

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-sematic and sparse-lexical features while beating the state-of-the-art methods on retrieval time. (Paper under review)

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. (Paper accepted at 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
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.

RECENT PUBLICATIONS

1. BLISS: A Billion scale Index using Iterative Re-partitioning, SIGKDD 2022
   Gaurav Gupta, Tharun Medini, Anshumali Shrivastava, and Alexander J. Smola
   Invited talk at NeurIPS 2021 BigANN competition track https://big-ann-benchmarks.com

2. 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

3. 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

IN REVIEW

1. CAPS: A Practical Partition Index for Filtered Similarity Search
   Gaurav Gupta, Jonah Yi, Benjamin Coleman, Chen Luo, Vihan Lakshman, Anshumali Shrivastava

2. IDentity with Locality: An ideal hash for efficient gene sequence search
   Gaurav Gupta, Tianyi Zhang*, Aditya Desai*, Anshumali Shrivastava*

3. HashOrder: Accelerating Graph Processing Through Hashing-based Reordering
   Tianyi Zhang, Aditya Desai, Gaurav Gupta, Anshumali Shrivastava

4. STORM: Foundations of End-to-End Empirical Risk Minimization on the Edge
   Gaurav Gupta, Ben Coleman, John Chen, Anshumali Shrivastava

RELEVANT COURSEWORK

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

Undergrad (Indian Institute of Technology, Hyderabad): Calculus, Complex variables and Linear Algebra , C/C++, Differential Equations, Linear Optimisation, Probability and Random Processes, Digital Signal Processing, Speech Signal Processing, Digital System Design, Image and Video Processing, Organisational Behaviour, Economics

INVITED TALKS/ PRESENTATIONS

BLISS- A Billion scale Index using Iterative Re-partitioning at KDD 2022
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.
RAMBO: Repeated And Merged BloOm Filter at
1. ACM SIGMOD 2021- International Conference on Management of data
2. Genome Informatics, 2020
3. Rice Oil & Gas High Performance Computing conference, 2020

POSITION OF RESPONSIBILITIES AND EXTRA-CURRICULARS

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)
8. Student Coordinator at Training and Placement Cell IIT Hyderabad. (2013-14)
9. Coordinator of Electronics Club IIT Hyderabad. (2012-13)
10. Participated at national level ABU ROBOCON- 2013 Pune. (2013)
11. Volunteer at National Service Scheme Team-IIT Hyderabad (2011)

OTHER ACHIEVEMENTS

1. Travel Grants – MLSys 2022 (NSF Funded), NeurIPS 2019.
2. Ph.D. panelist at undergrad student mentorship program at KDD 2022.
3. Nominated by the department 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.

GITHUB: https://github.com/gaurav16gupta

OTHER PUBLICATIONS

1. Siamese Networks for chromosomes classification, BIC, ICCV 2017 (pp. 72-81)
   Swati, Gaurav Gupta, M. Yadav, M. Sharma, L. Vig

2. Information Extraction from Hand-marked Industrial Inspection Sheets, ICDAR 2017 (vol 6 pp 33-38) IEEE
   Gaurav Gupta, Swati, M. Sharma, L. Vig

3. Indoor Localisation and Navigation on Augmented Reality Devices, IEEE ISMAR, 2016
   Gaurav Gupta, N. Kejriwal, P. Pallav, E. Hassan, S. Kumar, R. Hebbalaguppe

4. 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*

5. 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,