

GAURAV GUPTA

Electrical and Computer Engineering
Rice University, Houston, TX, United States

Email- gaurav.gupta@rice.edu
Webpage- gaurav16gupta.github.io
LinkedIn: [linkedin.com/in/gauravgupta16/](https://www.linkedin.com/in/gauravgupta16/)

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)

B.Tech, Major in Electrical Engineering

INTERESTS AND SKILLS

Hashing, Information Retrieval, Learning to hash/index, Machine Learning, Randomized Algorithms, Natural Language Processing, High Performance Computing, Computer Vision

Languages/Libraries: Python, C++, Pytorch, Tensorflow

PROFESSIONAL EXPERIENCE

Adobe Research, San Jose, Research Scientist Intern

(June 2022 - Aug2022)

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

Worked on data drift detection and continual learning on recommendation models.

Adobe Research, San Jose, Data Science Intern

(June 2021 - Aug2021)

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

Proposed a real-time product recommendation framework with both dense-semantic and sparse-lexical features, while beating the state-of-the-art methods on retrieval time. (Work under review)

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

(June 2020 - Aug2020)

Manager – *Dr. Alexander J. Smola (VP/Distinguished Scientist), Advisor- Prof. Anshumali Shrivastava*

Proposed a Billion scale near neighbour and extreme classification index with balanced load. (Paper accepted at SIGKDD 2022)

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

(May 2019 - Aug2019)

Manager - *Vijai Mohan (Senior Principal Applied Scientist), Advisor- Prof. Anshumali Shrivastava*

Worked on efficient sublinear and low memory indexing method 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. Worked on indoor localization systems and reconfigurable Inspection framework for augmented reality devices.

RESEARCH

- **BLISS:** BaLanced Index for Scalable Search. **(SIGKDD 2022)**
Billion scale fast near neighbour search using Learning to index
- **RAMBO:** Repeated And Merged BloOm filter **(SIGMOD 2021, Genome Informatics 2020)**
Terabyte scale ultrafast DNA search index
 - **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.
- **STORM:** Sketch Toward Online Risk Minimization **(In Review)**
Empirical Risk Minimisation using compressed sketches
- Constraint near neighbour search **(In Review)**
Single stage semantic and lexical match retrieval Index
- Cache Efficient Learning to Index. **(Ongoing research project)**
- Learned classifier and Index for zero-shot learning **(Ongoing research project)**
- Cache efficient Bloom Filter **(Ongoing research project)**
- Rank aggregation from pairwise comparison (MCMC sampling) **(PhD Qualifier)**

PUBLICATIONS/PREPRINTS

1. BLISS: A Billion scale Index using Iterative Re-partitioning, SIGKDD 2022
Gaurav Gupta, Tharun Medini, Anshumali Shrivastava, and Alexander J. Smola
2. STORM: Foundations of End-to-End Empirical Risk Minimization on the Edge (In review)
Gaurav Gupta, Ben Coleman, John Chen, Anshumali Shrivastava
3. 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
News Article: RAMBO speeds searches on huge DNA databases,
<https://www.sciencedaily.com/releases/2021/06/210628152920.htm>
4. 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
5. Siamese Networks for chromosomes classification, BIC, ICCV 2017 (pp. 72-81)
Swati, Gaurav Gupta, M. Yadav, M. Sharma, L. Vig
6. Information Extraction from Hand-marked Industrial Inspection Sheets, ICDAR 2017 (vol 6 pp 33-38) IEEE
Gaurav Gupta, Swati, M. Sharma, L. Vig
7. Indoor Localisation and Navigation on Augmented Reality Devices, IEEE ISMAR, 2016
Gaurav Gupta, N. Kejriwal, P. Pallav, E. Hassan, S. Kumar, R. Hebbalaguppe
8. 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
9. 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,

RELEVANT COURSEWORK

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

POSITION OF RESPONSIBILITIES AND EXTRA-CURRICULARS

- Mentor for 2 undergraduate summer interns at RUSH lab, Rice university (Summer 2022)
- Teaching assistant- Introduction to Deep Learning (Fall 2020)
(Prof. Ankit Patel, ELEC-COMP 516, ~80 students)
- Teaching assistant- Probabilistic Algorithms and Data-structures (Spring 2020)
(Prof. Anshumali Shrivastava, COMP 480/580, ~50 students)
- Reviewer at AISTAT-2021, CIKM-2022, NeurIPS-2022 (2021)
- Academic Mentor for Rice ECE Graduate Student Association (2020)
- Coordinator of Electronics Club IIT Hyderabad. (2012-13)

OTHER ACHIEVEMENTS

- NSF Student Travel Grant – MLSys 2022.
- PhD panellist at undergrad student mentorship program at KDD 2022.
- Ranked first at Neilsen’s Image Recognition Hackathon- 2017 for “Hole detection/ product out of stock” retail store challenge.
- Ranked 1913 (among top 0.4 %) in IIT-JEE engineering entrance exam, 2010.
- Selected for Deep Learning Summer School- MILA, Montreal, 2016.