# GAURAV GUPTA
Electrical and Computer Engineering
Rice University, Houston, TX, United States

Webpage- gaurav16gupta.github.io
Email- gaurav.gupta@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)**
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:** Python, C++, Pytorch, Tensorflow

## PROFESSIONAL EXPERIENCE

**Adobe Research, San Jose,** *Research Intern*      **(June 2021 - Aug2021)**
*Manager – Dr. Vishy Swaminathan (Sr. Principal Scientist)*
Proposed a real-time product recommendation framework with both dense-sematic 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 index with balanced load. Reduced 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) , 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), Advisor - Dr. Lovekesh Vig*
Worked on document reading, object recognition and classification for retail businesses, marker based head motion tracker. Other research projects on indoor localization systems for Google Glass and augmented reality based reconfigurable Inspection framework for head mounted devices. Learnt key technologies of image processing (motion tracking, text detection and object detection), convolutional neural networks, machine learning, reinforcement learning, algorithms and android app development.

**Uurmi Systems, Hyderabad** (Later acquired by MathWorks), *Research Intern*      **(June 2013 - July 2013)**
Worked on real time object tracking in aerial views.

## 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
- **STORM**: Sketch Toward Online Risk Minimization      **(In Review)**
  Empirical Risk Minimisation using compressed sketches
- Constraint near neighbour search      **(Ongoing research project)**
- Cache Efficient Learning to Index.      **(Ongoing research project)**
- Rank aggregation from pairwise comparison (MCMC sampling)      **(PhD Qualifier)**
- **Course Projects**
  - Learned Near Neighbor graph for Recommendation systems
  - A review on optimization methods for Learning to Hash
  - A review of Matrix Sensing
  - What2eat- Diet behaviour modelling using Transformer models
  - Python based simulation for Confocal Microscopy using Biobeam

## RELEVANT COURSEWORK

**Rice University, Houston TX:** Probabilistic Data structures, NLP seminar, Multi-Core Processing, Networks and Graph theory, Random Processes, Machine Learning, Optimization, Imaging through scattering medium, Statistical Signal Processing.

**Indian Institute of Technology, Hyderabad:** Calculus, Complex variables and Linear Algebra , C/C++ Programming Lab, Differential Equations, Linear Optimisation, Probability and Random Processes, Digital Signal Processing, Speech Signal Processing, Digital System Design, Computer Organisation, Image and Video Processing, Organisational Behaviour and Work Psychology, Economics

## INVITED TALKS/ PRESENTATIONS

**Presenting BLISS-** BaLanced Index for Scalable Search at ACM SIGKDD 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.

**Presented RAMBO:** Repeated And Merged BloOm Filter at
- ACM SIGMOD 2021- International Conference on Management of data
- Genome Informatics, 2020
- Rice Oil & Gas High Performance Computing conference, 2020

**Presented Document reader at**
- IAPR International Conference on Document Analysis and Recognition (ICDAR-2017), Kyoto, Japan.

**Presented AR Inspection Framework and Indoor Localisation and Navigation at**
- IEEE International Symposium on Mixed and Augmented Reality (ISMAR-2016), Mexico.

## 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_

4. 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. Siamese Networks for chromosomes classification,  BIC, ICCV 2017 (pp. 72-81)
   _Swati, Gaurav Gupta, M. Yadav, M. Sharma, L. Vig_

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

8. Indoor Localisation and Navigation on Augmented Reality Devices, IEEE ISMAR, 2016
   _Gaurav Gupta, N. Kejriwal, P. Pallav, E. Hassan, S. Kumar, R. Hebbalaguppe_

9. 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_

10. 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,_

## POSITION OF RESPONSIBILITIES AND EXTRA-CURRICULARS

- <u>Mentor</u> for 2 undergraduate summer interns at RUSH lab, Rice university (Summer 2022)
- <u>Teaching assistant</u>- Introduction to Deep Learning (Fall 2020)
  (Prof. Ankit Patel, ELEC-COMP 516, ~80 students)
- <u>Teaching assistant</u>- Probabilistic Algorithms and Data-structures (Spring 2020)
  (Prof. Anshumali Shrivastava, COMP 480/580, ~50 students)
- <u>Reviewer</u> at AISTAT-2021, CIKM-2022 (2021)
- Student <u>volunteer</u> at NeurIPS-2019, Vancouver, Canada (2019)
- Academic <u>Mentor</u> for Rice ECE Graduate Student Association (2020)
- <u>Media Head</u> (photography and videography) at Indian Students at Rice (ISAR) (2019-20)
- Student <u>Coordinator</u> at Training and Placement Cell IIT Hyderabad. (2013-14)
- <u>Coordinator</u> of Electronics Club IIT Hyderabad. (2012-13)
- Participated at national level ABU ROBOCON- 2013 Pune. (2013)
- <u>Volunteer</u> at National Service Scheme Team-IIT Hyderabad (2011)

## OTHER ACHIEVEMENTS

- NSF Student Travel Grant – MLSys 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