Gaurav Tarlok Kakkar

gaurav21776@gmail.com
404.731.2584
Portfolio

Research Interests

- **Systems for Machine Learning**
- **Databases**
- **Video Analytics**
- **Distributed Computing**

Education

- **Georgia Institute of Technology** — Atlanta, GA
  Master's in Computer Science; ***GPA: 4.0/4.0*** — *Aug. 2019 – ongoing*
- **Indian Institute of Technology** — Kanpur, India
  Bachelor of Engineering in Computer Science; ***GPA: 9.5/10.0*** — *Aug. 2013 – July. 2017*

Patents

- Text Wrap Detection GT Kakkar, M Singh - US Patent App. 16/190,562, 2020 (link)
- M. Rastogi, G.Kakkar, S. Sinha, P. Mehrotra. Mapping annotations to ranges of text across documents - US Patent App. 16/189,415 (link)
- M. Rastogi, G.Kakkar, S. Sinha, P. Mehrotra. Digital Annotation And Digital Content Linking Techniques - US Patent App. 16/372,037 (link)

Experience

- **Google** — Sunnyvale, USA
  Software Development Intern in Cloud SQL — *May 2020 - July 2020*
  - Lead the project to accelerate OLAP (Online Analytical Processing) queries by automatically building columnar cache indexes
  - Improved the query statistics collection engine and building ML driven columnar cache index advisor.
  - 5x improvement in query execution time with no manual cost overhead and filed US patent.
- **Adobe Systems India Pvt. Ltd.** — Noida, India
  Member of Technical Staff — *Jul 2017 - ongoing*
  - **Regenerate Layout from PDFs**: Research project
    - Built a multitude of ML algorithms processing together along with Software level heuristics to provide a single click layout generator from an inspiration pdf
    - Implemented a deep learning model, modified Faster RCNN to detect shape agnostic text wrap in a given pdf
    - Tackled challenges viz. detecting white space cover, creating master pages, organizing raw text runs into well defined text frames and intelligently figuring out object styles.

    Key Achievements: Filed patent in US on shape agnostic text wrap detection, implemented document analysis techniques and researched on core text properties.
  - **Import PDF Comments**: Keynote feature shipped with InDesign Max 2019
    - Implemented a novel approach to import and easily track the feedback made on a pdf version of document, solving the most in demand feature request of our million designers
    - Mastered the existing PDF library, to tackle the challenges of associating text or graphics with the annotation.

    Key Achievements: Filed patent in US and mastered PDF Library APIs
- **Adobe Systems India Pvt. Ltd.** — Bangalore, India
  Research Intern in Big Data Experience Lab — *May 2016 - July 2016*
  - Generating personalized bundles of products for customers of e-Commerce website that are driven by needs.
  - Formulated a novel approach of incorporated common sense knowledge, Concept net along with data driven insights.
  - Formed candidate set using hierarchical and minimum spanning tree based clustering algorithm over customer centric data enriched by semantic analysis. (slides)

Key Projects

- **[Machine Learning Systems] Exploratory Video Analytics** GaTech
  Research Project with Prof. Arulraj Joy *Jan. 2020 – Ongoing*
  - Implementing databases system for efficiently executing SQL like visual queries involving deep learning based UDFs.
  - Designed a Cascades style top-down rule-driven cost-based optimizer. Selectivity estimation of the UDFs, reuse aware plan enumeration and semantic-based reuse are the key components.
  - Collaborating with Georgia Tech's Agricultural Technology Research Program to monitor chicken behavior using camera feeds from poultry farms. (code)

- **[Machine learning Systems] Fast Array of Wimpy GPUs(FAWG)** GaTech
  Research Project with Prof. Alexey Tumanov *Sep. 2020 – Ongoing*
  - Serve memory-hungry models using model parallelism on cheap wimpy GPUs while meeting the latency SLOs
  - Proactive planner regresses over the batching parameters, model partitions, operator replicas, hardware types, and operator placement to search for a cost-effective model serving plan
  - The reactive planner behaves as a high-frequency tuner to auto-scale to meet tail latency goals in response to changes in the query arrival process.

- **[Computer Vision] Dense Image Captioning with NMS Convnet** IITK, India
  Research Project with Prof. Gaurav Sharma *Aug. 2016 – Nov 2016*
  - Analyzed the work - DenseCap by Andrej Karpathy et. al. - by experimenting with the parameters and design choices of Fully Convolutional Localization Network on Visual Genome dataset
  - To discard the existing test-time non-maximal suppression, used trainable spatial suppression layer from the work - A convnet for non-maximum suppression by Jan Hosang et. al. - to enhance the mAP of DenseCap from 5.698 to 5.76 (report) (slides)

- **[NLP] Automatic Abstract Generation for Research Papers** IITK, India
  Research Project with Prof. Harish Karnick *Aug. 2016 – Nov 2016*
  - Used a conglomerate of Extractive and Abstractive summarization techniques to generate abstracts for research/academic papers.
  - Used Topic Models, TextRank and Latent Semantic Analysis to extract important sentences which were fed into an RNN encoder-decoder network.
  - The model was trained on NIPS research papers and evaluated using the ROUGE metric. (report)

- **[Databases] Cafeteria Automation System** IITK, India
  Under-Graduate Project with Prof. Sumit Ganguly *Jan. 2016 – Dec 2016*
  - Designed a desktop app in C# incorporating mess menu creation, consumption Records, items BOM management, worker Attendance and salary management.
  - Won **Dr. Elizabeth & Varkey Cherian Award - Best UG project** with an impact on campus community.
  - As of May 2017, managed over 2,00,000 transactions of worth greater than INR 3.4 million. (slides) (code)

- **[CyberSecurity] Online Identity and Authentication using Blockchain** IITK, India
  Under-Graduate Project with Prof. Sandeep Shukla *Jan. 2017 – Apr 2017*
  - Design a novel protocol over Ethereum that leverages smart contracts and security of distributed ledger technology, to solve the problem of online identity and decentralized authentication while still preserving privacy.
  - Protocol covers all the key aspects viz. account creation(sign-up), account updation (authentication), account updation (authentication involving third party) and credential verification(sign-in). (report)(slides)

- **[Computer Architecture] Load Value Prediction** IITK, India
  Research Project with Prof. Mainak Chaudhuri *Jan. 2017 – Apr 2017*
  - Analyzed SPEC2006 benchmarks for the presence of Value Locality within instruction using different history depths.
  - Implemented a Load Value Prediction Unit to enable instruction level parallelism using PIN tool and analyzed its performance on the SPEC2006 benchmarks. (report)

- **[ML] Real Time Background Foreground Segmentation in Surveillance Video** IITK, India
  Research Project with Prof. Harish Karnick *Jan. 2016 – Apr 2016*
  - Classifying the input frames from campus surveillance video into foreground and background frames.

- Achieved an accuracy of 98.37% by implementing the **Codebook algorithm** augmented with the **Sobel-Feldman operator**. (report)

- **[Compiler] End to end Compiler for Go Programming Language** IITK, India
  Course Project with Prof. Subhajit Roy *Jan. 2016 – Apr 2016*
  - Implemented from scratch an entire pipeline of compiler; lexical analyzer, parser, semantic analyzer, conversion to intermediate language(MIPS), code generator.
  - Supported a wide range of Go Language features viz. arrays, nested loops, recursive function calls, type inference, arithmetic/boolean operations

- **[Operating Systems] NachOS** IITK, India
  Course Project with Prof. Mainak Chaudhari *Aug. 2015 – Nov 2015*
  - Extended the standard system call library for NachOS.
  - Implemented scheduling algorithms viz. UNIX scheduling, Round Robin, Non Preemptive and Shortest Job First.
  - Implemented page replacement algorithms viz. Random Page-Allocation, FIFO, LRO and LRU Clock.

Technical Skills

Languages	Python, C++, C, SQL, OpenMP, MPI, Js, Go, Bash, Assembly, HTML, CSS
ML	Tensorflow, Keras, scikit-learn, OpenCV

Achievements and awards

2017	Dr. Elizabeth & Varkey Cherian Award(Best UG project with an impact on campus community)
2017	Academic Excellence Award, IIT Kanpur (awarded to top 7% students in the institute)
2014	Academic Excellence Award, IIT Kanpur (awarded to top 7% students in the institute)
2013	All India Rank 236, IIT-JEE Advanced (among 150,000 candidates).

Extracurricular

- **Achievements**
  - Dr. Elizabeth & Varkey Cherian Award - Best UG project with an impact on campus community - 2017
  - Academic Excellence Award, IIT Kanpur - Equivalent to Dean's list in US - 2014, 2016
- **Conferences Attended**
  - Video analytics session at VLDB - Online - Sep, 2020
  - Poster Presentation at Adobe Tech Summit - San Francisco - Feb, 2019
  - Paper selected for Lightning Talk at Grace Hopper Celebration India - Nov, 2018
- **Mentoring and Leadership**
  - Mentor in Technovation - Drive to promote and motivate high school girls in the technology sector to improve gender equality
  - Graduate Teaching Assistant for Data Analysis using Deep Learning and Databases

Relevant Courses

- **Systems**: High Performance Parallel Computing, Advanced Operating Systems, Database Technologies, Computer Architecture, Compiler Design, Computer Networks, Computer Security, Advanced Data Structure and Algorithms
- **ML/Data Science**: Data analysis using Deep Learning, Recent Advances in Computer Vision, Machine Learning Tools and Techniques, Natural Language Processing, Data Visualization and Analysis