Ishank Sharma

Long Beach, California, Portfolio | LinkedIn | Github | Competitive Coding | Email: ishankdev@gmail.com | Phone: +1 (562)-341-2003

SUMMARY

I build high-performance, cost-efficient data systems that drive automation and business impact at scale. At CommerceIQ, a B2B unicorn SaaS company, I developed production-grade systems used by 2,200+ eCommerce brands, including Microsoft, Kellogg's, Nestlé, Colgate-Palmolive, and Johnson & Johnson, to optimize retail operations and increase profitability.

EDUCATION

California State University, Long Beach August 2024 - May 2026
Master of Science in Computer Science GPA: 4.0/4.0
Ramaiah Institute of Technology, Bengaluru August 2017 - May 2021
B.E in Information Science and Engineering GPA: 3.75/4.0

TECHNICAL SKILLS

- **Languages**: C, C++, Java, Python, Javascript
- **Libraries**: Spring , Node.js, Flask, Keras, Tensorflow, PyTorch, Scikit Learn
- **Databases and Cloud Technologies**: MYSQL, MongoDB, Redis, Databricks, Snowflake, Firebolt
- **Operating Systems**: Windows, Linux, MacOS
- **Software Types:** SaaS**,** Data Intensive Applications

EXPERIENCE

California State University, Long Beach – College of Business, Long Beach, California
Graduate Research Assistant *(Python, Deep Learning)* February 2025 - Present

- Identifying gaze patterns of the human eye on digital ads, potentially helping companies improve their click through rates and better targeting across demographics and age groups.

CommerceIQ, Bengaluru, India
Software Engineer II *(Python, Java SpringBoot , Databricks, SQL, Snowflake)* February 2021 - August 2024

- **Recommendation System:** Developed a budget recommendation pipeline using SQL, performing extensive data preprocessing and data cleaning on historical data, that predicted advertising spend with 95% accuracy, leading to a $200k reduction in wasted customer ad spend.
- **Designed and implemented** backend data infrastructure for a real-time advertising platform, encompassing data collection, data storage, and data processing of **500,000+ hourly ad data points**, ensuring data integrity and scalability.
- **Cold Start Solution:** Solved the cold start problem by designing a synthetic data strategy that blended sales and ad performance, enabling 90% accurate recommendations for clients with **zero historical data.**
- **Optimised Data Pipelines:** Reduced Snowflake pipeline costs by **70%** and improved resource allocation and resource optimisation through a strategy of date-specific backfilling, trimming resource use, and eliminating redundant processing, enhancing overall system planning.
- Engineered a custom SQL migration layer using extensible design patterns, streamlining database migration from Snowflake to Firebolt and reducing migration time by **40%.**
- **Migration from Snowflake to Databricks:** Orchestrated a full migration of the budget prediction feature from Snowflake to Databricks, slashing infrastructure costs by 52% and decreasing latency by **80% using indexing and broadcast joins.**

Internet Archive, San Francisco, California
Open Source Contributor March 2020 - December 2020

- **Archive.org:** Added keyboard support for zooming in/out on Safari and increased test coverage.
- **OpenLibrary.org:** Enhanced SEO by adding metadata to the index page, improving search visibility and ranking.
- **BookGenomeProject.org:** Developed an XML parser for identifying page types in Internet Archive books and introduced additional stop words for improved text cleanup.

LEADERSHIP

Apostle Incubator: Served as the sole tech consultant for Apostle Incubator, guiding MBA students.
Google Developers Student Club: Help with the day-to-day club activities and manage the club website.

ACADEMIC AND SIDE PROJECTS

- **Retrieval Augmented Generation Implementation AI**
  Constructed a retrieval-augmented generation pipeline utilising LangChain and Python, achieving a 40% reduction in AI hallucination while improving contextual accuracy based on BLEU scores.
- **Hyperparameter Tuning and Experiments with Deep Learning:** Investigated the impact of different activation functions on deep learning model performance and published findings to fix the three biggest causes of inaccurate image classification and gradient vanishing.

ACHIEVEMENTS

- **Winner**: Smart India Hackathon
- **Runner-Up**: Microsoft IncubateIND Hackathon
- **Runner Up:** Reverie Language Hackathon