Asoke Datta

☎ +1(209)777-9694 ✉ adcseus1@gmail.com • <u>LinkedIn</u> • <u>GitHub</u> • <u>Google Scholar</u>

EDUCATION

- **Ph.D. Candidate** in Computer Science | UC Merced | Fall 2018 - Summer 2024(Expected)
- **BSc** in Computer Science and Engineering | Leading University, Sylhet, Bangladesh | 2010 - 2013

COURSES

- Algorithm Design and Analysis • Computer Architecture • Database System Implementation • Parallel Computing • Operating Systems
- Advanced Topic in Intelligent Systems • Big Data Science • Data Structure • Distributed Systems • Computer Networks • Compilers

TOOLS

- C • CPP • Python • JAVA • Javascript • Bash • GSQL • SQL • CUDA • PostgreSQL • MonetDB • Oracle • DuckDB • SQLite
- HEAVY.AI • TigerGraph • PyTorch • docker • gprof • git • GDB • Linux • VMWare • Hyper-V

EXPERIENCE

- **Research Assistant(Database Systems) | University of California, Merced |** Aug 2018 – Present
  - Benchmark database query optimizers. Analyze and document performance trade-offs.
  - Design, develop, evaluate, and document new query optimization algorithms.
  - Design experimental methodology.
  - Tools: [*CPP*, *Python*, *Bash*, *SQL*], [*PostgreSQL*, *SQLite*, *Oracle, DuckDB, MonetDB, HEAVY.AI*], docker, git
  - Publications:
    a. Yesdaulet Izenov, <u>Asoke Datta</u>, Jun Hyung Shin, Florin Rusu. COMPASS: Online Sketch-based Query Optimization for In-Memory Databases. <u>Sigmod 2021</u>
    b. <u>Asoke Datta</u>, Yesdaulet Izenov, Brian Tsan, Florin Rusu. Simpli-Squared: Optimizing Without Cardinality Estimates. Link: <u>arxiv.org</u> (SiMod 2024 - Sigmod Affiliated Workshop)
    c. <u>Asoke Datta</u>, Brian Tsan, Yesdaulet Izenov, Florin Rusu. Analyzing Query Optimizer Performance in the Presence and Absence of Cardinality Estimates. Link: <u>arxiv.org</u>
    d. Brian Tsan, <u>Asoke Datta</u>, Yesdaulet Izenov, Florin Rusu. Approximate Sketches. <u>Sigmod 2024</u>
    e. Yesdaulet Izenov, <u>Asoke Datta</u>, Brian Tsan, Florin Rusu. Sub-optimal Join Order Identification with L1-error. <u>Sigmod 2024</u>

- **Teaching Assistant | University of California, Merced** | Aug 2018 - Present
  - Conduct guest lectures, labs, and discussion sessions.
  - Supervise design and development of student class projects.
  - Evaluate Student Performance and share feedback.
  - Tools: [*Python, Java*], [*SQLite*], git

- **Ph.D. intern, Database Query Optimization | TigerGraph |** May 2022 – Aug 2022
  - Benchmarking graph database using Social Network Benchmark (SNB) suite.
  - Generate synthetic data for the graph database. Controlling the distribution of data.
  - Evaluate histogram estimation quality.
  - Tools: [*Python, Javascript, GSQL*], [*TigerGraph*], git

- **System Engineer | Accenture**, **Bangladesh** | Oct 2014 – Nov 2017
  - Deploy and manage physical and virtual server environments.
  - Develop methodologies for the automation of manual operations.
  - Problem troubleshooting, service delivery as per SLA, and documentation of major events.
  - Tools: [*Bash, Python*], VMWare, Hyper-V, Linux, and Windows servers

PROJECTS

- **Database Implementation** | Spring 2019
  - Objective**:** Implement database main components including a) Catalog, b) Query Optimizer, c) Data Loader, and d) Execution Engine.
  - Tools**:** [C, *CPP, Lex, YACC*], git, gprof, gdb; Repo: <u>github.com</u>
  - Result**:** Full working database pipeline (syntax limited).

- **Cardinality Estimation** | Spring 2021
  - Objective**:** Estimating the Cardinality of a database query using sampling, histogram, sketches, and ML(CNN) model
  - Tools**:** *Python*; Repo: <u>github.com</u>
  - Result**:** Successfully estimated the cardinality of a database query.

- **ML Projects (SOFC Approximation, Autonomous Retail)** | Fall 2019, Spring 2020
  - Objective: Approximate and optimize Solid Oxide Fuel Cell simulation.Event Detection, Object Recognition for Autonomous Retail.
  - Tools: [*Python, CPP*], TensorFlow, OpenFuelCell, Cantera, PostNet; Repo: <u>github.com</u>
  - Result**:** SOFC - Model accuracy 78 percent on synthetic data
    Autonomous Retail - Partial implementation; Image classification task: 70 percent accuracy.

OTHER ACTIVITIES

- Sub-reviewer [14 reviews]: ACM SIGMOD, ACM DEBS, IEEE, SSDBM.
- Attended ACM ICPC national and inter-university contests from 2010 - 2013.