Asoke Datta

☎ +1(209)777-9694 ✉ adatta2@ucmerced.edu • linkedin.com/in/ad26 • github.com/Asoke26

EDUCATION

- Ph.D. Candidate in Computer Science **|** UC Merced **|** 2018 - Present
- BSc in Computer Science and Engineering **|** Leading University, Sylhet **|** 2010 - 2013

COURSES

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

TECHNICAL SKILLS

- CPP • Python • PostgreSQL • MonetDB • MapD • Oracle • DB2 • SQL • Machine Learning • Tensorflow • CUDA • JAVA • bash/shell
- Javascript • AWS • gprof • git • GDB • docker • Machine Learning(ML) • GSQL • TigerGraph

EXPERIENCE

- **Research Assistant |** University of California, Merced **|** Aug 2018 – Present
  - Working on finding efficient techniques to optimize database queries.
  - Developed models and scripts to generate synthetic workload and manipulate benchmark data based on experimental needs.
  - Currently working on understanding the correlation between DB query execution and optimization.
- **Teaching Assistant (Database Systems) |** University of California, Merced **|** Aug 2018 - Present
  - Conducted guest lectures and labs in class of Max. 120 students.
  - Supervise design and development of student class projects.
  - Evaluate Student Performance and share feedback.
- **Ph.D. intern, Query Optimization |** TigerGraph **|** May 2022 – Aug 2022
  - Benchmarking graph database.
  - Generate synthetic data for the graph database. Controlling the distribution of data.
  - Evaluate histogram estimation quality.
- **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.

PROFESSIONAL ACTIVITIES

- Sub-reviewer, Scientific and Statistical Database Management Conference [6 reviews]
- Sub-reviewer, IEEE Transactions on Knowledge and Data Engineering [4 reviews]
- External-reviewer, IEEE International Conference on Big Data [2 reviews]
- Sub-reviewer, ACM SIGMOD Internation Conference on Management of Data [1 review]
- Sub-reviewer, ACM International Conference on Distributed and Event-based Systems [1 review]

PUBLICATIONS

- Yesdaulet Izenov, Asoke Datta, Jun Hyung Shin, Florin Rusu. COMPASS: Online Sketch-based Query Optimization for In-Memory Databases. Sigmod 2021, Link: dl.acm.org/doi/abs/10.1145/3448016.3452840
- Yesdaulet Izenov, Asoke Datta, Jun Hyung Shin, Florin Rusu. Online Sketch-based Query Optimization. Link: arxiv.org/abs/2102.02440
- Asoke Datta, Yesdaulet Izenov, Brian Tsan, Florin Rusu. Simpli-Squared: A Very Simple Yet Unexpectedly Powerful Join Ordering Algorithm Without Cardinality Estimates. Link: arxiv.org/abs/2111.00163

PROJECTS

- **Database Implementation |** Spring 2019
  - Objective**:** Implement database main components including a) Catalog, b) Query Optimizer, c) Data Loader, and d) Execution Engine.
  - Tools**:** CPP, Lex, YACC; Repo: github.com/Asoke26/Database_Implementation
  - Result**:** Full working database pipeline (syntax limited); input: query; Output: result
- **Cardinality Estimation |** Spring 2021
  - Objective**:** Estimating the Cardinality of a database query using sampling, histogram, sketches, and ML(CNN) model
  - Tools**:** Python; Repo: github.com/Asoke26/Cardinality-Estimation
  - Result**:** Programs estimated 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.; Repo: github.com/Asoke26/OpenFuelCell
  - Result**:** SOFC - Model accuracy 78 percent on synthetic data, runtime optimized by 98 percent
    Autonomous Retail - Partial implementation; PostNet accuracy 90 percent (20 cases); Image Classifier 70 percent accuracy.