Sami Hadouaj

+1 313 414 0110 | shadouaj@umich.edu | linkedin.com/in/sami-hadouaj/ | Michigan,USA

Summary

I am a Ph.D. candidate at the University of Michigan-Dearborn. My field of research is about building the bridge between database theory and machine learning algorithms. I possess comprehensive expertise in system programming, advanced C++ development, and proficiency in parallel computing, optimization, and performance optimization methodologies.

Education

University of Michigan (Rackham Graduate School) Dearborn, Michigan
Doctor of Philosophy in Computer and Information Science Jan 2022 – Jun 2026

- Relevant coursework: Database systems , Text mining and information retrieval , Algorithms analysis and design

Tunisia Polytechnic School Tunis, Tunisia
Master of Science in Computer Science Sep 2020 – Jun 2021

National Institute of Applied Science and Technology Tunis, Tunisia
Bachelor in Software Engineering Sep 2016 – Jun 2021

Skills

Languages: C/C++, Stan, Java, Python, SQL

Technologies: OpenMP, MPI, Apache Arrow, ClangJIT, Perf, Cmake, Ninja, Maven, Docker, Git, Valgrind, Linux

Libraries: Scikit-Learn, NumPy, Pandas, Matplotlib, Seaborn, Eigen C++

Areas: Databases, Machine learning, Probabilistic Programming, Parallel Computing, Bayesian Inference, Optimization

Work Experience

Graduate Research Assistant Jan 2022 – Present
University of Michigan-Dearborn Michigan, USA

- Conducted research merging probabilistic programming and databases, resulting in a unified framework optimized for highly efficient statistical inference based on MCMC sampling.
- Contributed to the creation of a query engine capable of encoding probabilistic programs and performing inference tasks. This resulted in a highly optimized query engine written in C++ and leveraging Apache Arrow, Just in time compilation (ClanJIT), and parallel MCMC inference (OpenMP).
- Implemented parallelization within our probabilistic query execution engine, including the optimization of the inference algorithm and effective resolution of race conditions. This strategic enhancement led to a substantial x8 increase in overall system performance.
- Worked on algorithm optimization. This was done by running performance monitoring tools to identify the algorithm's behavior at each step and make decisions on how to optimize it. An example application would be the implementation of a stochastic collapsed variational inference algorithm that was 50% faster than the original implementation.

Software Engineering Intern Feb 2021 – Jun 2021
Dearborn AI Research Center (University Of Michigan-Dearborn) Michigan, USA

- Incorporated artificial intelligence within the realm of software engineering with the goal of significantly reducing developer overhead. This strategic adoption of AI aimed at enhancing efficiency and productivity.
- Reduced development overhead by creating a chatbot capable of addressing open issues on GitHub and Jira pertaining to software quality. The chatbot successfully improved the quality metrics used to quantify software development.

Research Intern Jul 2020 – Sep 2020
Dearborn AI Research Center (University Of Michigan-Dearborn) Michigan, USA

- Boosted development process efficiency by creating an automated system responsible for generating refactoring commit messages.
- Applied BERT embeddings and attention mechanisms for the autonomous generation of descriptive GitHub commit messages This provided significant assistance to developers throughout the development process.

Publications

Ouael Ben Amara *, **Sami Hadouaj ***, Niccolo Meneghetti. (*SIGMOD 2024*) "StarfishDB: A Query Execution Engine For Relational probabilistic Programming."

Talks & presentations

SIGMOD 2024 Conference: Talk and presentation of StarfishDB in the main research track.
(Santiago, Chile June 2024)

North East Database Day: Talk and presentation about query driven inference in probabilistic databases.
(Boston, MA March 2023)

Certificates

Machine learning specialization, Coursera.

Deep learning specialization, Coursera.