Sami Hadouaj

shadouaj@umich.edu — samihadouaj.github.io — linkedin.com/in/sami-hadouaj — github.com/samihadouaj

Summary

I am a PhD candidate (All But Dissertation) in the Computer and Information Science department at the University of Michigan-Dearborn, advised by Dr. Niccolò Meneghetti. My primary research goal is to integrate Bayesian machine learning into database systems to natively manage uncertain and probabilistic data. I am proficient in C++, with extensive experience in approximate inference, probabilistic programming, and systems development. I am passionate about tackling complex problems at the intersection of database systems and machine learning.

Education

University of Michigan (Rackham Graduate School) — Dearborn, Michigan
Ph.D in Computer and Information Science — GPA: 3.96/4 — Advised by: Dr. Niccolò Meneghetti — Jan 2022 – 2026

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

Research Experience

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

- Designed a system that automatically generates optimized variational inference algorithms for probabilistic Datalog models, enabling scalable Bayesian inference through Stochastic Variational Inference.
- Accelerated inference performance by applying knowledge compilation techniques to reduce the number of variational parameters, improving computational efficiency while preserving model accuracy.
- Engineered a high-performance general-purpose probabilistic programming system in C++. The system integrates Apache Arrow for efficient in-memory data handling, ClangJIT for just-in-time compilation of inference operations, and OpenMP for parallelized execution. It achieved performance competitive with specialized, model-specific algorithms in benchmark evaluations.

Graduate Research Assistant — May 2020 – May 2021
University of Michigan-Dearborn — Dearborn, MI

- Developed an AI chatbot integrated with GitHub and Jira that leveraged natural language processing to analyze software quality issues and recommend solutions to developers, significantly enhancing development productivity. Advised by Dr. Marouane Kessentini.

Publications

- Stochastic variational inference for Datalog probabilistic programming **[Under review]**
  ***Sami Hadouaj**, Ouael Ben Amara, Niccolò Meneghetti*
- StarfishDB: a query execution engine for relational probabilistic programming* **[SIGMOD 2024]**
  Ouael Ben Amara, **Sami Hadouaj***, Niccolò Meneghetti*
  [DOI] [Code] [Poster] [Slides]
- Reproducibility report for ACM SIGMOD 2024 paper: "On efficient large sparse matrix chain multiplication" **[SIGMOD 2024 Reproducibility]**
  ***Sami Hadouaj**, Xiaojun Dong*
  [DOI]
- Reproducibility report for ACM SIGMOD 2024 paper: "StarfishDB: a query execution engine for relational probabilistic programming" **[SIGMOD 2024 Reproducibility]**
  ***Sami Hadouaj**, Ouael Ben Amara, Niccolò Meneghetti, Alexander Krause, Georgiy Lebedev*
  [DOI]

Skills

Programming Languages: C++, Python, SQL, Java, Datalog, Bash

Systems & Tools: Linux, Git, Docker, CMake, LLVM/Clang (ClangJIT), Apache Arrow, Ninja, gdb/Valgrind, perf, OpenMP

Data & ML: NumPy, Pandas, scikit-learn, Matplotlib, Eigen

Research Areas: Database Systems, Approximate Bayesian Inference, Machine Learning, Probabilistic Programming, Probabilistic Circuits

Languages: English (Fluent), French (Intermediate)

Service

BIGDATA (2022, 2023, 2024, 2025): Reviewer

ICDM (2025): Reviewer

SIGMOD ARI (2024): Artifacts reproducibility, evaluation, and availability reviewer

Volunteered with Accelerate4kids to create educational materials and resources that support learning opportunities for underprivileged children. (Sep - Dec 2024)

Teaching

Lab Instructor, CIS 200: Software Engineering — FA 2025

Teaching Assistant, CIS 556: Database Systems (with Dr. Niccolò Meneghetti) — FA 2023, WI 2024, FA 2024, WI 2025

Teaching Assistant, CIS 4951: Senior Design I and II (with Dr. Bruce Maxim) — SU 2025

Honors & Awards

Rackham Doctoral Intern Fellowship Program: Selected and funded by the University of Michigan Rackham Graduate School to conduct research during Summer 2024.

Talks & Presentations

Talks

SIGMOD 2024 Conference: *StarfishDB: A Query Execution Engine for Relational Probabilistic Programming* Santiago, Chile · June 2024

Northeast Database Day: *Query-driven Inference in Probabilistic Databases* Boston, MA · March 2023

Posters & Presentations

Northeast Database Day: *Declarative Probabilistic Programming for Explainable AI* Boston, MA · January 2025

SIGMOD 2024 Conference: *StarfishDB: A Query Execution Engine for Relational Probabilistic Programming* Santiago, Chile · June 2024

Northeast Database Day: *StarfishDB: A Query Execution Engine for Relational Probabilistic Programming* Boston, MA · March 2024

Northeast Database Day: *Query-driven Inference in Probabilistic Databases* Boston, MA · March 2023

Graduate Courses

Computer Science: Database Systems, Text Mining and Information Retrieval, Algorithm Analysis and Design

Mathematics: Probability and Statistics

References

- Niccolò Meneghetti
  Assistant Professor, Department of Computer and Information Science, University of Michigan-Dearborn
  Email address: niccolom@umich.edu

Hobbies

Running; Cycling; Rock Climbing