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, Michigan

- Designed an automated framework that generates optimized Stochastic Variational Inference algorithms for probabilistic Datalog models, enabling scalable Bayesian 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++, integrating Apache Arrow for efficient in-memory data management, ClangJIT for just-in-time compilation of inference operations, and OpenMP for parallel execution. Achieved performance comparable to specialized, model-specific algorithms in benchmark evaluations.

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

- 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

- **Sami Hadouaj**, Ouael Ben Amara, Niccolò Meneghetti. "Stochastic variational inference for Datalog probabilistic programming" *(Under review)*

- Ouael Ben Amara[*], **Sami Hadouaj**[*] , Niccolò Meneghetti. "StarfishDB: a query execution engine for relational probabilistic programming" [1] *Proceedings of the ACM on Management of Data* 2.3 (2024): 1-31. doi:10.1145/3654988 (**SIGMOD 2024**)
  [Paper] [Code] [Poster] [Slides]

- **Sami Hadouaj**, Xiaojun Dong. "Reproducibility report for ACM SIGMOD 2024 paper: 'On efficient large sparse matrix chain multiplication' " *Reproducibility Reports of the 2024 International Conference on Management of Data.* 2025. doi:10.1145/3687998.3717039 (**SIGMOD 2024 Reproducibility**)

---

[1]Sami Hadouaj and Ouael Ben Amara are both first authors of this paper.

- Krause, Alexander, Georgiy Lebedev, **Sami Hadouaj**, Ouael Ben Amara, and Niccolò Meneghetti "Reproducibility report for ACM SIGMOD 2024 paper: 'StarfishDB: a query execution engine for relational probabilistic programming' " *Reproducibility Reports of the 2024 International Conference on Management of Data.* 2025. doi:10.1145/3687998.3717056 (**SIGMOD 2024 Reproducibility**)

Skills

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

Systems & Tools: Linux, Git, Docker, CMake, Apache Arrow, Ninja, gdb/Valgrind, perf, OpenMP

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

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 reviewer: Conducted independent code evaluation and experiment reproduction.

Volunteered with Accelerate4kids, developing educational materials to expand learning access 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

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

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

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

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

Graduate Courses

Database Systems (CIS 556), Text Mining and Information Retrieval (CIS 536), Algorithm Analysis and Design (CIS 505), Data Analytics in Software Engineering (CIS 580), Mathematical Statistics (MATH 525)

Hobbies

Running; Cycling; Rock Climbing

References

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

- Qiang Zhu
  Professor, Department of Computer and Information Science, University of Michigan-Dearborn
  Email address: qzhu@umich.edu

- Srijita Das
  Assistant Professor, Department of Computer and Information Science, University of Michigan-Dearborn
  Email address: sridas@umich.edu