

Sina Sheikholeslami

Ph.D. Student,
Division of Software and Computer Systems
School of Electrical Engineering and Computer Science
KTH Royal Institute of Technology, Stockholm, Sweden

sinash@kth.se
<https://kth.se/profile/sinash>
<https://linkedin.com/in/sinasheikholeslami>

Summary

I'm a soon-to-graduate Computer Science Ph.D. student specializing in computer systems, distributed deep learning (DL), data-intensive computing, machine learning (ML), and applications of machine learning and deep learning. I have an M.Sc. degree in data science and a minor in innovation and entrepreneurship, which I have been awarded by KTH and TU Eindhoven. Before starting my graduate studies, I was part of the Big Data R&D Engineering team at Digikala.com, the largest e-commerce business and the most-visited website in Iran, where I was a major contributor to the establishment of their first Big Data Analytics stack. We reported directly to both the CTO and the CEO of the company, and there, I collaborated with several teams in the organization, such as Marketing, Warehouse Logistics, Network Infrastructure, and Quality Assurance, in designing and delivering products that would address their needs. I have a solid foundation and experience in data-intensive computing, data science, and ML/DL engineering, as evident through my experiences both in the industry and academia. I have also participated in many cross-disciplinary projects, and **have published papers on applications of machine learning, deep learning, and data mining in earth observation and remote sensing, public transit systems, e-commerce, and wood microscopy.**

I have been the Head Teaching Assistant of two well-known courses at KTH, ID2221 Data-intensive Computing and ID2222 Data Mining, which hundreds of M.Sc. students at KTH and elsewhere have taken over the past 5 years and who are now successful engineers in the Swedish tech sector. I have also supervised 17 M.Sc. degree projects. My award-winning Ph.D. research has made an impact in both academia and industry, mainly through providing production-ready frameworks and approaches for distributed, efficient, large-scale training of deep neural networks.

Education

Ph.D. in ICT, specialization in Software and Computer Systems
KTH Royal Institute of Technology

Defended on 2025-03-27

M.Sc. in Data Science (minor in Innovation and Entrepreneurship)
KTH Royal Institute of Technology & Eindhoven University of Technology

2019

B.Sc. in Software Engineering,
Amirkabir University of Technology (Tehran Polytechnic)

2017

Interests

Distributed deep learning, Data-parallel deep learning, Systems for machine learning, Data-intensive computing, Applications of machine learning and deep learning, Artificial Intelligence for research

Professional Experiences

Doctoral Student, KTH Royal Institute of Technology

September 2019 - Present

- During my PhD studies, I have worked on systems for large-scale and distributed machine learning and deep learning. In particular, we developed and released the first framework for automated, parallel ablation studies for deep learning, called "AutoAblation", as part of Maggy; introduced a novel approach for dataset partitioning in data-parallel training of deep neural networks by considering the importance of dataset examples, named Importance-aware DPT, which won the Best Artifact Award of DAIS 2023, and a novel approach to weight initialization of deep neural networks using model weights from the hyperparameter tuning trials. I have also been the academic supervisor of 17 master thesis degree projects.

Master Thesis Research Intern, Hopsworks AB and RISE SICS

January 2019 - August 2019

- I worked on my M.Sc. Thesis as a Research Intern at the Computer Systems Laboratory (CSL) of RISE SICS as part of Hopsworks AB. Our work resulted in "Maggy", a framework for asynchronous distributed machine learning built on top of Apache Spark.

Big Data R&D Engineer, Digikala.com

September 2016 - August 2017

- Digikala.com is the leading e-commerce business and the number one most-visited website in Iran. As a member of the R&D team, I contributed to a number of projects, including the Clickstream project for collecting, storing and analyzing user activity data, and the LMAP project in which we developed and implemented an enterprise-wide error monitoring and log analysis solution. I also contributed to the design and maintenance of our Big Data Analytics infrastructure. I worked with many technologies and products in development and production settings, including Apache Spark, Apache Kafka, Elasticsearch, Snowplow, and Splunk. I left Digikala after almost a year to pursue my graduate studies.

Tools and Skills

Python, PyTorch, LangGraph, Apache Spark, Ray, TensorFlow, Apache Kafka, HDP Stack, Elastic Stack, Hopsworks, Amazon Web Services, Microsoft Azure, Google Cloud Platform, SQL, Splunk

Awards

Best Artifact Award, The 23rd IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS), 2023.

EIT Digital Merit Scholarship, EIT Digital Academy, 2017-2019.

Selected Publications

For a complete list of publications please see **my Google Scholar profile**.

Selected Refereed Conference Articles in Chronological Order

1. **Sina Sheikholeslami**, Hamid Ghasemirahni, Amir H. Payberah, Tianze Wang, Jim Dowling, and Vladimir Vlassov. "Utilizing Large Language Models for Ablation Studies in Machine Learning and Deep Learning." In Proceedings of the 5th Workshop on Machine Learning and Systems, co-located with EuroSys, 2025.
2. **Sina Sheikholeslami**, Tianze Wang, Amir H. Payberah, Jim Dowling, and Vladimir Vlassov. "Deep Neural Network Weight Initialization from Hyperparameter Tuning Trials." In Proceedings of the 31st International Conference on Neural Information Processing (ICONIP), 2024.
Code: <https://github.com/ssheikholeslami/dnn-weight-initialization-from-hp-tuning/>

3. Tobias Johannesson, Isak Rubensson, **Sina Sheikholeslami**, Ahmad Al-Shishtawy, and Vladimir Vlassov. "DUGET: Leveraging Machine Learning for Dynamic User Grouping and Evolution Tracking in Public Transit Systems." In Proceedings of the IEEE International Conference on Big Data (Big Data), pp. 1785-1794. 2024.
4. Zihao Xu, Peter Nordström, **Sina Sheikholeslami**, Ahmad Al-Shishtawy, and Vladimir Vlassov. "A Semi-Supervised Model for Non-Cellular Elements Segmentation in Microscopy Images of Wood." In Proceedings of the IEEE International Conference on Big Data (Big Data), pp. 2049-2056. 2024.
5. **Sina Sheikholeslami**, Amir H. Payberah, Tianze Wang, Jim Dowling, and Vladimir Vlassov. "The Impact of Importance-Aware Dataset Partitioning on Data-Parallel Training of Deep Neural Networks." In Proceedings of the 23rd IFIP International Conference on Distributed Applications and Interoperable Systems (DAIS), pp. 74-89. 2023.
Best Artifact Award
Code: <https://github.com/ssheikholeslami/importance-aware-data-parallel-training>
6. Albert Asratyan, **Sina Sheikholeslami**, and Vladimir Vlassov. "A Parallel Chain Mail Approach for Scalable Spatial Data Interpolation." In Proceedings of the IEEE International Conference on Big Data (Big Data), pp. 306-314. 2021.
7. **Sina Sheikholeslami**, Moritz Meister, Tianze Wang, Amir H. Payberah, Vladimir Vlassov, and Jim Dowling. "AutoAblation: Automated Parallel Ablation Studies for Deep Learning." In Proceedings of the 1st Workshop on Machine Learning and Systems, co-located with EuroSys, pp. 55-61. 2021.
Code: <https://github.com/ssheikholeslami/ablation-paper-experiments>
8. Marina Angelovska, **Sina Sheikholeslami**, Bas Dunn, and Amir H. Payberah. "Siamese Neural Networks for Detecting Complementary Products." In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop, pp. 65-70. 2021.
9. Moritz Meister, **Sina Sheikholeslami**, Amir H. Payberah, Vladimir Vlassov, and Jim Dowling. "Maggy: Scalable Asynchronous Parallel Hyperparameter Search." In Proceedings of the 1st Workshop on Distributed Machine Learning, co-located with CoNEXT, pp. 28-33. 2020.
10. Moritz Meister, **Sina Sheikholeslami**, Robin Andersson, Alexandru A. Ormenisan, and Jim Dowling. "Towards Distribution Transparency for Supervised ML with Oblivious Training Functions." In Workshop on MLOps Systems, co-located with MLSys. 2020.

Journal Articles

11. Desta Haileselassie Hagos, Theofilos Kakantousis, **Sina Sheikholeslami**, Tianze Wang, Vladimir Vlassov, Amir Hossein Payberah, Moritz Meister, Robin Andersson, and Jim Dowling. "Scalable Artificial Intelligence for Earth Observation Data Using Hopsworks." Remote Sensing 14, no. 8. 2022.
12. Desta Haileselassie Hagos, Theofilos Kakantousis, Vladimir Vlassov, **Sina Sheikholeslami**, Tianze Wang, Jim Dowling, Claudia Paris et al. "ExtremeEarth Meets Satellite Data from Space." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 14. 2021.

Theses

13. **Sina Sheikholeslami**. "Ablation Programming for Machine Learning." M.Sc. Thesis, KTH. 2019.

Academic Services

Program Committee Member

AutoML, International Conference on Automated Machine Learning, 2025
ACL-SRW, ACL Student Research Workshop, 2019, 2022, 2023, 2024, 2025
EACL-SRW, EACL Student Research Workshop, 2021, 2023, 2024
ACL-IJCNLP-SRW, ACL-IJCNLP Student Research Workshop, 2021
NAACL-SRW, NAACL Student Research Workshop, 2021, 2025

External Reviewer

Euro-Par, International European Conference on Parallel and Distributed Computing, 2025
JSys, The Journal of Systems Research, 2023
CLOUD, IEEE International Conference on Cloud Computing, 2022
UCC, IEEE/ACM International Conference on Utility and Cloud Computing, 2021

Reproducibility Reviewer

JSys Artifact Evaluation Board Member, 2024 -
AutoML, International Conference on Automated Machine Learning, 2023, 2025

Selected Talks

Parallel Ablation Studies for Machine Learning with Maggy on Apache Spark, Data + AI Summit Europe, 2020
Asynchronous Hyperparameter Tuning and Ablation Studies with Apache Spark, CASTOR Software Days, 2019
Scalable Deep Learning for Copernicus Data, ESA Earth Observation Phi-Week, 2019

Supervision and Teaching Experiences

KTH Royal Institute of Technology, Master Thesis Supervisor

For a complete list of the theses I have supervised please see **my KTH profile**.

KTH Royal Institute of Technology, Teaching Assistant

ID2221 Data Intensive Computing (2019 - 2024)

ID2222 Data Mining (2019 - 2023)

ID1206 Operating Systems (2022 - 2023)

Eindhoven University of Technology, Teaching Assistant

2IAB0 Data Analytics for Engineers, 2017

Tehran Polytechnic, Teaching Assistant

Operating Systems, 2014

Principles of Computer and Programming, 2012

References

Vladimir Vlassov, Professor, Div. of Software and Computer Systems, KTH

`vladv@kth.se`

Amir H. Payberah, Associate Professor, Div. of Software and Computer Systems, KTH

`payberah@kth.se`

Jim Dowling, Co-founder & CEO, Hopsworks

`jim@hopsworks.ai`

April 10, 2025