Rush Tabesh

Ph.D. Student, ISTA

🛿 Vienna, Austria 🔹 🗹 stabesh@ista.ac.at 🔹 🕲 tabesh.me 🔹 in Rush Tabesh 🔹 🗘 RushTabesh

Research Interests

- Efficient Deep Learning
- Quantization-Aware Training (QAT)
- Neural Network Compression
- Parameter-Efficient Fine-Tuning (PEFT)

Education

Institute of Science and Technology Austria (ISTA) Ph.D. Student in Computer Science (Machine Learning) Advisor: Prof. Dan Alistarh

Sharif University of Technology B.Sc. in Computer Science

Klosterneuburg, Austria Sep. 2024 - present

> Tehran, Iran Sep. 2019 - Jul. 2024

Publications

([†] denotes equal contribution)

S. Ashkboos[†], M. Nikdan[†], S. Tabesh[†], R. L. Castro, T. Hoefler, D. Alistarh, *HALO: Hadamard-Assisted Lower-Precision Optimization for LLMs. arXiv pre-print.*

A. Panferov, J. Chen, **S. Tabesh**, R. L. Castro, M. Nikdan, D. Alistarh, *QuEST: Stable Training of LLMs with* 1-Bit Weights and Activations. In Workshop on Sparsity in Large Language Models (SLLM), ICLR 2025.

E. Zverev, S. Abdelnabi, S. Tabesh, M. Fritz, C. H. Lampert, *Can LLMs Separate Instructions From Data? And What Do We Even Mean By That?* In International Conference on Learning Representations (ICLR), 2025.

E. Zverev, E. Kortukov, A. Panfilov, A. Volkova, **S. Tabesh**, S. Lapuschkin, W. Samek, C. H. Lampert, *ASIDE: Architectural Separation of Instructions and Data in Language Models*. In *Workshop on Building Trust in LMs*, ICLR 2025.

M. Nikdan[†], **S. Tabesh**[†], E. Crnčević, D. Alistarh, *RoSA: Accurate Parameter-Efficient Fine-Tuning via Robust Adaptation*. In *International Conference on Machine Learning (ICML)*, 2024.

D. Kuznedelev[†], **S. Tabesh**[†], K. Noorbakhsh[†], E. Frantar[†], S. Beery, E. Kurtic, D. Alistarh, *TACO: Vision Models Can Be Efficiently Specialized via Few-Shot Task-Aware Compression. Transactions on Machine Learning Research (TMLR)*, in press 2025.

Selected Research Experiences

Distributed Algorithms and Systems Laboratory, ISTA Research Intern (in-person) — Supervisor: Prof. Dan Alistarh Klosterneuburg, Austria Sep 2022 – Oct 2023

Tehran – Iran

Jun 2023 – Jul 2024

- RoSA: Robust Adaptation of Large Language Models Proposed a parameter-efficient fine-tuning framework based on sparse/low-rank RPCA-style adaptations, improving LLM downstream accuracy while keeping training memory and compute overhead minimal. z
- TACO: Task-Aware Compression for Vision Models Extracted specialised subnetworks from large vision backbones (ViT, ConvNeXt, ...) via layer-wise, data-aware pruning and distillation, achieving 2×-5× inference speed-ups with negligible task-specific accuracy loss.

Sharif University of Technology

Bachelor Thesis — Supervisor: Prof. M. H. Yassaee

Optimal Transport for Knowledge Distillation in Pruned Vision Models — Applied Wasserstein loss to
mitigate training instability of sparsified networks, enabling reliable knowledge transfer and improved robustness
versus KL-based distillation.

Graduate Courses

- Foundations of Theoretical Deep Learning (Ph.D. Course)
- High Dimensional Statistics (Course Project: Margin-Based Generalization Lower Bounds)
- Introductory Natural Language Processing (Course Project: Multi Modal Sentiment Analysis)
- Statistical Methods and Time Series in Finance

Selected Teaching Experiences

Teaching Assistant, Sharif University of Technology ([†]: *Graduate Course*) Deep Learning[†] • Introduction to Statistics • Data Structures and Algorithm Design Advanced Programming **Selected Working Experience** Data Scientist @ Divar: Semantic retrieval of real estate ads; Geo-search. Mar. 2023 – Jul. 2023 ML Engineer @ Hezardastan: Image retrieval of mathematical text. Mar. 2022 – Dec. 2022 Skills Technical Skills: Python | C/C++ | CUDA | Julia | Java | SLURM Machine Learning Frameworks: PyTorch (Advanced) | Jax | Ray **Language Skills:** Persian (*native*) | English (C2) | German (A1) **Extra Curricular Activities** Problem setter, Codeforces 2018 Educational Codeforces Round 57 Math Teacher/Mentor, Rastaa (non-profit edu. NGO) Tehran 2019 - 2020 References **Prof. Dan Alistarh** dan.alistarh@ista.ac.at Professor, Institute of Science and Technology Austria (ISTA) Prof. Mohammad Hossein Yassaee yassaee@sharif.edu

Assistant Professor, EE Department, Sharif University of Technology

Please refer to my personal website for more information

- Differential Privacy (Course Project: Differentially Private Matrix Factorization)
- Computer Vision
- Deep Learning
- Econometrics
- Stochastic Processes