Rush Tabesh

Ph.D. Student, ISTA, Austria

stabesh@ista.ac.at • tabesh.me • Rush Tabesh • GitHub • Google Scholar

Research Interests

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

Education

Institute of Science and Technology Austria (ISTA) — Klosterneuburg, Austria
Ph.D. Student in Computer Science (Machine Learning) — *Sep. 2024 - present*
Advisor: Prof. Dan Alistarh

Sharif University of Technology — Tehran, Iran
B.Sc. in Computer Science — *Sep. 2019 - Jul. 2024*

Publications

(† **denotes equal contribution**)

S. Tabesh†, M. Safaryan†, A. Panferov, A. Volkova, D. Alistarh, *CAGE: Curvature-Aware Gradient Estimation For Accurate Quantization-Aware Training*. *arXiv pre-print*.

R. L. Castro, A. Panferov, **S. Tabesh**, O. Sieberling, J. Chen, M. Nikdan, S. Ashkboos, D. Alistarh, *Quartet: Native FP4 Training Can Be Optimal for Large Language Models*. In *Advances in Neural Information Processing Systems* (*NeurIPS*), 2025.

S. Tabesh†, S. Ashkboos†, M. Nikdan†, R. L. Castro, T. Hoefler, D. Alistarh, *HALO: Hadamard-Assisted Lower-Precision Optimization for LLMs*. In *Advances in Neural Information Processing Systems* (*NeurIPS*), 2025.

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 *International Conference on Machine Learning* (*ICML*), 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.

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

S. Tabesh†, D. Kuznedelev†, 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*).

Review Experience

NeurIPS 2025 Reviewer (*Top Reviewer*)

ICLR 2025 Reviewer

Other Research Experiences

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

- **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.
- **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 — Tehran – Iran

Bachelor Thesis — Supervisor: Prof. M. H. Yassaee — *Jun 2023 – Jul 2024*

- **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.

Selected Courses and Course Projects

Graduate Courses

- Foundations of Theoretical Deep Learning
- 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
- Differential Privacy (Course Project: Differentially Private Matrix Factorization)
- Modern Machine Learning
- Numerical Methods and Algorithms
- Stochastic Processes

Selected Teaching Experiences

Teaching Assistant, Sharif University of Technology

(†: *Graduate Course*)

- Deep Learning†
- Data Structures and Algorithm Design
- Introduction to Statistics
- Advanced Programming

Work 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 – Sep. 2022*

Software Engineer @ Hezardastan: Back-end Engineer — *Mar. 2021 – Mar. 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 and Mentor, Rastaa (non-profit edu.) 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