muradtuk@gmail.com

My LinkedIn profile

Morad Tukan

Tel: +972543298724

Machine/Deep learning │ Big Data │ Approximation algorithms │ Robotics

Birth: Haifa, 1995

Bio

I am a Computer Science Ph.D. graduate from the University of Haifa, specializing in machine learning, deep learning, big data, and robotics. My doctoral research was conducted under the guidance of Prof. Dan Feldman at the Robotics and Big Data Labs and Prof. Roee Diamant at the Underwater Acoustic and Navigation Lab.

Throughout my academic journey, I have honed my skills in developing practical solutions for real-world challenges. My work focuses on enhancing state-of-the-art solvers through approximation algorithms and computational geometry. I am passionate about bridging the gap between theoretical and applied computer science, driving innovation, and delivering impactful results.

Education

Ph.D. Graduate, Computer Science, University of Haifa, Haifa, Israel　　March'18 - September'22
Thesis: Coreset construction frameworks for machine learning models
Advisor: Dan Feldman and Roee Diamant

M.sC. Graduate, Computer Science, University of Haifa, Haifa, Israel (Magna Cum Laude)　　October'14 - May'17
Thesis: Coreset construction framework for convex functions
Advisor: Dan Feldman

B.sC. Graduate, Etgar program, Computer Science, University of Haifa, Haifa, Israel　　October'10 - August'14
Etgar program – BsC degree in Computer Science for excellent high-school students

Relevant Experience

Lead Researcher　　May'22 - Present
DataHeroes

- The goal of the team is to boost the training of ML models and DL neural networks through novel subset selection techniques, incurring the smallest degradation in performance.
- We have provided the first provable subset selection techniques for deep learning neural networks.
- We created the first general framework for subset selection given an ML model, and its loss function, allowing other researchers and practitioners to curate their data with ease.
- Developed the first provable subset selection for handling SQL count queries, for tabular categorical data.

Graduate Research Assistant　　October'15 - September'22
University of Haifa, Prof. Dan Feldman

- Developed novel sampling-based data summarization approaches and frameworks to accelerate training of ML models, reducing memory storage and running time.
- Research in provable compression methods and neural network model optimization.
- Developed novel provable compression of the search space of sampling-based path planners.

Graduate Research Assistant   March'21 - September'22
University of Haifa, Prof. Roee Diamant

- Developed a novel water current interpolation system for marine science datasets via data compression techniques, graph theory and clustering techniques.

Visiting Graduate Research Student   May'17 - September'17
MIT Computer Science & Artificial Intelligence Laboratory, Prof. Daniela Rus

- Investigated novel sampling-based approach with provable guarantees for accelerating SVMs.

Teaching assistant   February'16 - July'16
University of Haifa   November'21 - April'22

- Introduction to Computer science
- Processing of Marine Databases

Tutor for young students in the fields of computer science and mathematics   September'16 – April'17
Leo Baeck Education Center Haifa Israel (Israel national service)   November'17 – March'18

# Highlighted Publications (* stands for equal contribution)

[1] M. Tukan, L. Mualem, and M. Feldman. "Practical 0.385-Approximation for Submodular Maximization Subject to a Cardinality Constraint". In: *Advances in Neural Information Processing Systems*. Ed. by A. Globerson, L. Mackey, D. Belgrave, A. Fan, U. Paquet, J. Tomczak, and C. Zhang. Vol. 37. Curran Associates, Inc., 2024, 51223–51253 (conference rank: A^*).

[2] L. Mualem, M. Tukan, and M. Feldman. "Bridging the Gap between General and Down-Closed Convex Sets in Submodular Maximization". In: *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, IJCAI-24*. Ed. by K. Larson. Main Track (conference rank: A^*). International Joint Conferences on Artificial Intelligence Organization, Aug. 2024, pp. 1926–1934. DOI: `10.24963/ijcai.2024/213`.

[3] M. Tukan, E. Biton, and R. Diamant. "An efficient drifters deployment strategy to evaluate water current velocity fields". In: *IEEE Journal of Oceanic Engineering* (2024, (Q1 Journal, Impact factor: 3.883)).

[4] A. Maalouf*, M. Tukan*, N. Loo, R. Hasani, M. Lechner, and D. Rus. "On the Size and Approximation Error of Distilled Datasets". In: *Thirty-seventh Conference on Neural Information Processing Systems*. 2023, (conference rank: A⋆).

[5] M. Tukan, S. Zhou, A. Maalouf, D. Rus, V. Braverman, and D. Feldman. "Provable Data Subset Selection For Efficient Neural Networks Training". In: *Proceedings of the 40th International Conference on Machine Learning*. Ed. by A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett. Vol. 202. Proceedings of Machine Learning Research. PMLR, 23–29 Jul 2023, 34533–34555 (conference rank: A⋆).

[6] A. Maalouf*, M. Tukan*, V. Braverman, and D. Rus. "AutoCoreset: An Automatic Practical Coreset Construction Framework". In: *Proceedings of the 40th International Conference on Machine Learning*. Ed. by A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett. Vol. 202. Proceedings of Machine Learning Research. PMLR, 23–29 Jul 2023, 23451–23466 (conference rank: A⋆).

[7] M. Tukan*, L. Mualem*, and A. Maalouf*. "Pruning neural networks via coresets and convex geometry: Towards no assumptions". In: *Advances in Neural Information Processing Systems* 35 (2022), 38003–38019 (conference rank: A⋆).

[8] M. Tukan, A. Maalouf, D. Feldman, and R. Poranne. "Obstacle aware sampling for path planning". In: *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2022, 13676–13683 (conference rank: A).

[9] M. Tukan*, A. Maalouf*, M. Weksler*, and D. Feldman. "No Fine-Tuning, No Cry: Robust SVD for Compressing Deep Networks". In: *Sensors* 21.16 (2021), 5599 (Journal rank: Q1, Impact factor: 3.576).

[10] M. Tukan, C. Baykal, D. Feldman, and D. Rus. "On coresets for support vector machines". In: *Theoretical Computer Science* 890 (2021), 171–191 (Journal rank: Q2, Impact factor: 0.827).

[11] M. Tukan, A. Maalouf, and D. Feldman. "Coresets for near-convex functions". In: *Advances in Neural Information Processing Systems* 33 (2020 (Conference rank: A⋆)).

Selected Projects

- *Generative AI*
  - **Data curation for text-to-image (T2I) diffusion model (on Slurm)**
    Sampling based data curation on Text-to-Image datasets to accelerate the fine-tuning and refinement of Diffusion models.
  - **LLM data curation (on Slurm)**
    High-quality subset sampling of text datasets for accelerating the fine-tuning of various LLM models and removing redundancy in the training data.

- *DL model compression*
  - **NLP deep neural networks structural pruning (collaborated with Samsung)**
    Created a framework for compressing NLP deep neural networks using ellipsoidal matrix sketching techniques.
  - **General framework of neuron and filter pruning**
    A unified framework of for pruning neurons in MLP Layers and filters of CNN layers with minimal needed constraints.

- *Applications of classical machine learning*
  - **Spark-based SVM acceleration (collaborated with ShieldIOT)**
    Developed a spark-based framework for anomaly detection while adapting support vector machines to streaming settings through provable subset selection techniques (namely coresets).
  - **Device identification (collaborated with ShieldIOT)**
    A spark based framework for device identification using data summarization techniques for projective clustering in steaming settings.

Selected Talks

TAMC 2020 (Virtual) │ NeurIPS 2020 (Virtual) │ NeuripsIL 2020 (Virtual) │ Robotics and Big Data (RBD) labs - University of Haifa 2021 │ Departmental seminar Marine Technologies - University of Haifa 2022 │ AISTATS 2022 (Virtual) │ IROS 2022 (Kyoto, Japan) │ NeurIPS 2022 (New Orleans, USA) │ ICML 2023 (Honolulu, Hawaii) │ NeurIPS 2023 (New Orleans, USA) │ IJCAI 2024 (Jeju Island, South Korea) │ NeurIPS 2024 (Vancouver, Canada)

Honors

Excellence scholarship for PhD candidates │A scholarship for outstanding students (the Municipality of Haifa) │CRI Hackaton (6th)

Program Committee (PC) Member / Reviewer

- IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI) 2020
- Conference on Neural Information Processing Systems (NeurIPS) 2021
- International Conference on Learning Representations (ICLR) 2021
- International Conference on Artificial Intelligence and Statistics (AISTATS) 2021
- International Conference on Machine Learning (ICML) 2022
- Transactions on Machine Learning Research (TMLR) 2022
- Conference on Neural Information Processing Systems (NeurIPS) 2022
- Conference on Neural Information Processing Systems (NeurIPS) 2023
- Transactions on Machine Learning Research (TMLR) 2023
- International Conference on Learning Representations (ICLR) 2023
- Transactions on Machine Learning Research (TMLR) 2024
- Conference on Neural Information Processing Systems (NeurIPS) 2024

- IEEE Journal of Oceanic Engineering (IEEE JOE) 2024

- International Conference on Machine Learning (ICML) 2025

- Transactions on Machine Learning Research (TMLR) 2025

- Conference on Neural Information Processing Systems (NeurIPS) 2025

Skills & Interests

Software:	Python, PyTorch, Scikit-learn, Numpy, Matlab, CVXPy, SQL, JGraphT, PySpark, Slurm, C++, C, Java, YALMIP
Languages:	English (fluent), Hebrew (fluent), Arabic (native)
Interests:	Traveling, sports (swimming, hiking), math riddles, riddle solving