REZA SHIRKAVAND

Department of Computer Science, University of Maryland, College Park, MD, USA
Email, Website, Google Scholar, Github

RESEARCH INTERESTS

Efficient GenAI, Machine Learning

EDUCATION

Sharif University of Technology *2020*
B.Sc in Computer Engineering

University of Maryland *2022 - Present*
Ph.D. in Computer Science
Research Focus: Efficient Gen , Mixture of Experts

SELECTED PUBLICATIONS

Cost-Aware Contrastive Routing for LLMs
R. Shirkavand, *S. Gao, P. Yu, H. Huang*
NeurIPS 2025 (Spotlight)

Bilevel ZOFO: Bridging Parameter-Efficient and Zeroth-Order Techniques for Efficient LLM Fine-Tuning and Meta-Training
R. Shirkavand, *P. Yu, Q. He, H. Huang*
NeurIPS 2025

Efficient Fine-Tuning and Concept Suppression for Pruned Diffusion Models
R. Shirkavand, *P. Yu, S. Gao, G. Somepalli, T. Goldstein, H. Huang*
CVPR 2025

Not All Prompts Are Made Equal: Prompt-based Pruning of Text-to-Image Diffusion Models
R. Shirkavand*, *A. Ganjdanesh*, S. Gao, H. Huang*
ICLR 2025

ARGUS: Hallucination and Omission Evaluation in Video-LLMs
R. Rawal, ***R. Shirkavand***, *H. Huang, G. Somepalli, T. Goldstein*
ICCV 2025

ToMoE: Converting Dense Large Language Models to Mixture-of-Experts through Dynamic Structural Pruning
S. Gao, T. Hua, ***R. Shirkavand***, *C. Lin, et. al*
Under Review

Pruning Without Fine-Tuning: Dynamic Pruning of Autoregressive Image Generation Models to Mixtures of Experts
R. Shirkavand, *S. Gao, H. Huang*
Under Review

From Pixels to Prose: A Large Dataset of Dense Image Captions
V. Singla, K.Yue, ***R. Shirkavand***, *S. Paul, et al.*
Preprint

Incomplete Multimodal Learning for Complex Brain Disorders Prediction
***R. Shirkavand**, L. Zhan, H. Huang, L. Shen, P.M. Thompson*
Preprint

WORK EXPERIENCE

PhD Research Intern Jun 2025 - Present
Roblox

Working on Mixture of Experts for Recommendation and Generation.

ML Researcher/Engineer Feb 2020 - Dec 2021
Netbina

Developed sentiment analysis models to help crisis management for clients, created topic detection models to identify emerging trends in news articles and tweets related to various industries, built a new multi-class classification model using convolutional neural networks, and implemented object detection models to increase the speed and accuracy of images processing.

TEACHING EXPERIENCE

Co-Instructor, Advanced Machine Learning Topics *Spring 2024*
Presented an overview of the transformer architecture and its applications.

Teaching Assistant, Systems & Projects Engineering *Summer 2022*
Taught essential topics in systems engineering and project management.

Teaching Assistant, Algorithmic Thinking *Spring 2022*
Delivered lectures and supported students in developing the theoretical and practical skills necessary for designing algorithms.

Teaching Assistant, Machine Learning *Fall 2019*
Supported students in gaining theoretical and practical knowledge in machine learning and statistical pattern recognition.

Teaching Assistant, Computer Networks *Spring 2019*
Facilitated the design and development of network applications and the implementation of conventional network management and routing protocols for students.

Teaching Assistant, Operating Systems *Fall 2018*
Guided and supported students in understanding and implementing the central concepts of operating systems through the development of a real, working, and simple kernel.

HONORS AND AWARDS

Nationwide University Entrance Exam - Mathematics *2015*
Ranked 14th among 181000 participants

Nationwide University Entrance Exam - Foreign Languages *2015*
Ranked 15th among 7000 participants

SKILLS

Programming Languages	Python, Java, C++, Matlab
Libraries	Pytorch, Tensorflow
Languages	English (Fluent) , French (Basic), Persian (Native)