Kuat Gazizov

Merced, CA, United States
kgazizov@ucmerced.edu linkedin.com/in/kgazizov Personal Website

Education

University of California, Merced **Jan 2022 – present**
PhD in Computer Science *Merced, CA, United States*
Supervisor: Miguel Á. Carreira-Perpiñán

Lomonosov Moscow State University **Sep 2015 – Jun 2019**
Bachelor of Science in Mathematics *Moscow, Russia*

Publications

[**NeurIPS 2025**] K. Gazizov and M. Á. Carreira-Perpiñán: "A faster training algorithm for regression trees with linear leaves, and an analysis of its complexity". *Advances in Neural Information Processing Systems, 2025* (**In press**)

[**BayLearn 2025 (Oral)**] K. Gazizov and Y. Idelbayev and M. Á. Carreira-Perpiñán: "Emerging Aspects in ResNet Quantization with Adaptive Codebook Sizes". *Extended abstract at Bay Area Machine Learning Symposium, 2025*

[**NeurIPS 2024**] M. Á. Carreira-Perpiñán and K. Gazizov: "The tree autoencoder model, with application to hierarchical data visualization". *Advances in Neural Information Processing Systems, 2024*

[**BayLearn 2024**] M. Á. Carreira-Perpiñán and K. Gazizov: "Hierarchical data visualization via PCA trees". *Extended abstract at Bay Area Machine Learning Symposium, 2024*

K. Gazizov and A. Zharmagambetov and M. Á. Carreira-Perpiñán: "A critical comparison of soft vs hard oblique classification trees". (**In submission**)

[**BayLearn 2023**] K. Gazizov and A. Zharmagambetov and M. Á. Carreira-Perpiñán: "Pros and cons of soft vs hard decision trees". *Extended abstract at Bay Area Machine Learning Symposium, 2023*

Projects

Mixed-bit Quantization using LC Algorithm | *ResNet Compression, Learning-Compression (LC), Adaptive Quantization*

- Developed a mixed-bit quantization method using the Learning-Compression (LC) algorithm. Formulated quantization as a constrained optimization problem to adaptively select per-layer bit-widths and remove redundant layers. Applied to ResNet(20-110) architectures, achieving up to 25x compression ratio with no accuracy loss.

Experience

University of California, Merced **Jan 2022 – present**
Research/Teaching Assistant *Merced, CA, United States*

- Research area: neural network compression (adaptive mix-bit quantization), learning oblique trees, tree-based methods for dimensionality reduction, clustering and regression problems.
- Assisting in teaching the following courses: Introduction to Machine Learning (Fall 2025), Data Structures (Fall 2025), Algorithm Design and Analysis (Spring 2025), Advanced Programming (Fall 2024), Introduction to OOP (Spring 2023), Introduction to programming (Fall 2022), Discrete Mathematics (Spring 2022)

Snap Inc. **May 2025 – Aug 2025**
Generative ML Intern *Santa Monica, CA, United States*

- Designed a lightweight module for a GAN, creating a general-purpose style (text prompt) transfer model.
- Replaced a multi-hour, single-style training pipeline; the new model provides instant zero-shot results and allows for optional, style-specific fine-tuning in minutes.

Sberbank **Jun 2020 – Oct 2021**
NLP ML Engineer *Moscow, Russia*

- Participated in the development and fine-tuning of a NLP framework for Named Entity Recognition (NER) and classification.

Professional Activities

Reviewer: Journal of Machine Learning Research (JMLR 2023), Neural Information Processing Systems (NeurIPS 2024, 2025), ACM SIGKDD (KDD 2024), International Conference on Learning Representations (ICLR 2024)

Technical Skills

Languages: Python (PyTorch, scikit-learn, numpy, scipy, pandas, transformers, LightGBM, XGBoost), C/C++,
Technologies: Linux, Git, Matlab