Max VLADYMYROV, PhD

Senior Research Scientist at Google DeepMind

Summary

The primary focus of my research is on improving efficiency, interpretability, reasoning, and adaptation capabilities of LLMs. Currently, my interests lie in understanding and utilizing in-context learning to improve model reasoning and adaptation. This builds on my expertise in meta-learning, nonlinear optimization, and manifold learning. My work has resulted in over 25 publications in top ML venues (NeurIPS, ICML, ICLR, CVPR) and impactful product launches across Google.

Personal data

Phone: +1 415 707 9685
E-mail: vladymyrov@gmail.com
Website: max-vladymyrov.github.io

Work Experience

Aug. 2020 to present: .... Senior Research Scientist, Google DeepMind

Co-led a team of 6 RS & SWE. Authored 15+ research papers primarily focuses on efficiency and understanding of LLMs. Recent contributions include:

- Developed a method for effective in-context compression in LLMs, enabling efficient processing of long context data (preprint).
- Proposed a practical meta-optimizer achieving up to 5x speedup for LLM fine-tuning (preprint).
- Formalized a connection between in-context learning and gradient descent (ICML'23 oral, NeurIPS'24).
- Developed methods for on-the-fly weight generation using Transformers and HyperNetworks (ICML'22, TMLR'24).

Apr. 2016 to Aug. 2020: .... Senior Research Scientist, Google Research

- Developed sequence-based models (Transformers, BERT, Mixture of Experts) to predict user satisfaction from user interaction logs.
- Designed models for addressing presentation bias in search.
- Applied these models to enhance ranking and triggering mechanisms in Google Search, ultimately handling close to 100% of the search traffic.

Sep. 2014 to Apr. 2016: .... Research Scientist, Yahoo! Labs

- Developed semantic analysis and a query triage system using embedding clustering techniques.
- Built the company's first deep learning search ranking algorithm.
- Improved performance of the email spam filtering system.

Aug. 2009 to Aug. 2014: .... Graduate Research Assistant, UC Merced

- Published 8 papers (7 as a first author) in leading ML conferences on novel techniques for efficient large-scale nonlinear optimization.
- Created efficient manifold learning algorithms for high-quality visualization of large datasets (1M+ points) using standard CPUs.

EDUCATION

Aug. 2009 to Aug. 2014: . . . . PhD, Computer Science
University of California, Merced. GPA: 3.9/4.0
Thesis: Large-Scale Methods for Nonlinear Manifold Learning.

Sep. 2007 to July 2008: . . . . MS, Computer Science (summa cum laude)
Kharkiv National University, Ukraine. GPA: 4.0/4.0
Thesis: Design and Optimization of Cargo Transportation Delivery.

Sep. 2003 to July 2007: . . . . BS, Applied Mathematics (summa cum laude)
Kharkiv National University, Ukraine. GPA: 3.9/4.0
Thesis: Stability of Some Stochastic Models in Economics.

Sep. 2004 to July 2009: . . . . BS and MS, Intern. Econ. Relations (cum laude)
Kharkiv National University, Ukraine. GPA: 3.7/4.0
Thesis: Crisis of the Global Financial System and Its Impact on the Ukrainian Economy.

PUBLICATION LIST, GOOGLE SCHOLAR

🧠 In-Context Learning, Mechanistic Interpretability, Algorithm Discovery

▷ Aleksandar Petrov, Mark Sandler, Andrey Zhmoginov, Nolan Miller, **Max Vladymyrov**: "Long Context In-Context Compression by Getting to the Gist of Gisting", *preprint*, 2025.

▷ Chen Sun, Renat Aksitov, Andrey Zhmoginov, Nolan Miller, **Max Vladymyrov**, Ulrich Rueckert, Been Kim, Mark Sandler: "How new data permeates LLM knowledge and how to dilute it", *ICLR 2025*, **spotlight**.

▷ **Max Vladymyrov**, Johannes von Oswald, Mark Sandler, Rong Ge: "Linear transformers are versatile in-context learners", *NeurIPS 2024*.

▷ **Max Vladymyrov**, Johannes von Oswald, Nolan Miller, Mark Sandler: "Efficient linear system solver with transformers", *AI for Math Workshop, ICML 2024*.

▷ Chen Sun, Nolan Miller, Andrey Zhmoginov, **Max Vladymyrov**, Mark Sandler: "Learning and unlearning of fabricated knowledge in language models", *Mechanistic Interpretability Workshop, ICML 2024*.

▷ Andrey Zhmoginov, Jihwan Lee, **Max Vladymyrov**, Mark Sandler: "Learning fast and slow: representations for in-context weight modulation", *Workshop on In-Context Learning, ICML 2024*.

▷ Johannes von Oswald, Eyvind Niklasson, Maximilian Schlegel, Seijin Kobayashi, Nicolas Zucchet, Nino Scherrer, Nolan Miller, Mark Sandler, **Max Vladymyrov**, Razvan Pascanu, João Sacramento: "Uncovering mesa-optimization algorithms in transformers", *preprint*, 2023.

▷ Johannes von Oswald, Eyvind Niklasson, Ettore Randazzo, João Sacramento, Alexander Mordvintsev, Andrey Zhmoginov, **Max Vladymyrov**: "Transformers learn in-context by gradient descent", *ICML 2023*, **oral**.

🌱 Few-Shot Learning, HyperNetworks

▷ **Max Vladymyrov**, Andrey Zhmoginov, Mark Sandler: "Continual HyperTransformer: a meta-learner for continual few-shot learning", *TMLR 2024*.

▷ Andrey Zhmoginov, Mark Sandler, **Max Vladymyrov**: "HyperTransformer: model generation for supervised and semi-supervised few-shot learning", *ICML 2022.*

🔧 Nonlinear Optimization, Model Adaptation

▷ Mark Sandler, Andrey Zhmoginov, **Max Vladymyrov**, Nolan Miller: "Training trajectories, mini-batch losses and the curious role of the learning rate", *preprint*, 2023.

▷ Utku Evci, Bart van Merrienboer, Thomas Unterthiner, **Max Vladymyrov**, Fabian Pedregosa: "GradMax: growing neural networks using gradient information", *ICLR 2022.*

▷ Mark Sandler, Andrey Zhmoginov, **Max Vladymyrov**, Andrew Jackson: "Fine-tuning image transformers using learnable memory", *CVPR 2022.*

▷ Alexander D'Amour, ..., **Max Vladymyrov**, ...: "Underspecification presents challenges for credibility in modern machine learning", *JMLR 2022.*

🎓 Meta-Learning, Learning-to-Learn

▷ Gus Kristiansen, Mark Sandler, Andrey Zhmoginov, Nolan Andrew Miller, Anirudh Goyal, Jihwan Lee, **Max Vladymyrov**: "Narrowing the focus: learned optimizers for pretrained models", *preprint*, 2024.

▷ Andrey Zhmoginov, Mark Sandler, Nolan Miller, Gus Kristiansen, **Max Vladymyrov**: "Decentralized learning with multi-headed distillation", *CVPR 2023.*

▷ Mark Sandler, **Max Vladymyrov**, Andrey Zhmoginov, Nolan Miller, Tom Madams, Andrew Jackson, Blaise Agüera Y Arcas: "Meta-learning bidirectional update rules", *ICML 2021.*

🗺 Manifold Learning, Dimensionality Reduction

▷ **Max Vladymyrov**: "No pressure! Addressing the problem of local minima in manifold learning algorithms", *NeurIPS 2019.*

▷ **Max Vladymyrov** and M. Á. Carreira-Perpiñan: "Fast, accurate spectral clustering using locally linear landmarks" *IJCNN 2017.*

▷ **Max Vladymyrov** and M.Á. Carreira-Perpiñán: "The variational Nyström method for large-scale spectral problems". *ICML 2016.*

▷ M.Á. Carreira-Perpiñán and **Max Vladymyrov**: "A fast, universal algorithm to learn parametric nonlinear embeddings". *NeurIPS 2015.*

▷ **Max Vladymyrov** and M.Á. Carreira-Perpiñán: "Linear-time training of nonlinear low-dimensional embeddings", *AISTATS 2014.*

▷ **Max Vladymyrov** and M.Á. Carreira-Perpiñán: "Locally linear landmarks for large-scale manifold learning". *ECML 2013.*

▷ **Max Vladymyrov** and M.Á. Carreira-Perpiñán: "Entropic affinities: properties and efficient numerical computation". *ICML 2013.*

▷ **Max Vladymyrov** and M.Á. Carreira-Perpiñán: "Partial-Hessian strategies for fast learning of nonlinear embeddings". *ICML 2012.*

Professional service

▷ Co-orgnized ICLR 2023 Workshop "From Cells to Societies: Collective Learning Across Scales".

▷ JMLR Editorial Board *(2020–present)*.

▷ Area chair: *NeurIPS (2025)*.

▷ Conference reviewer: *ICML (2018–2024), NeurIPS (2013–2025), ICLR (2021), ICCV (2017), ECCV (2016), WWW (2015–2016), SDM (2013), KDD (2015–2016), IJCAI-ML (2015)*.

▷ Top reviewer award: *NeurIPS (2020, 2024)*.

▷ Journal reviewer: *PAMI, Connection Science, Neurocomputing, Pattern Recognition, Pattern Recognition Letters, IJNS, IEEE Transactions on Neural Networks and Learning Systems*.