Katrina Drozdov (Evtimova)

kevtimova.github.io | kve216@nyu.edu | kevtimova | d4xi2HIAAAAJ | Katrina Drozdov

Summary

Interested in applied research on post-training techniques for large language models, such as stabilizing training by increasing sample diversity and mitigating reward hacking. My PhD introduced regularization methods that prevent collapse and enhance representation diversity in self-supervised systems. Recently, I fine-tuned Qwen2.5-7B using GRPO to study reasoning behavior and analyze failure modes that reduce pass@k and downstream performance. I aim to translate these insights into more reliable and broadly useful foundation models.

Education

Ph.D. in Data Science, New York University — Sep 2018 - Sep 2024
Thesis: "Representation Learning with Regularized Energy-Based Models".
Committee: Yann LeCun (Advisor), Kyunghyun Cho, Carlos Fernandez-Granda, Brenden Lake, Leon Bottou.

M.Sc. in Data Science, New York University — Sep 2015 - May 2017
Mentors: Kyunghyun Cho, David Sontag, Yacine Jernite.

B.A. in Mathematics, Harvard College — Sep 2009 - May 2013

Experience

LLM Evaluation Contributor, Snorkel AI — Feb 2025 - May 2025
Created and validated challenging reasoning tasks to probe and enhance LLM performance.

Research Intern, Meta - Fundamental AI Research (FAIR) — May 2020 - Aug 2020
Applied variance regularization to prevent collapse in sparse visual encoders, improving robustness and transfer.

Research Intern, Meta - Fundamental AI Research (FAIR) — May 2019 - Aug 2019
Developed hierarchical sparse representation learning methods for images.

Research Engineer, eBay - Recommendations — Jul 2017 - Aug 2018
Built machine learning models deployed in production for large-scale item recommendations.

Research Assistant, New York University - CILVR Lab — Oct 2016 - Mar 2017
Implemented Markov Logic Networks for clinical data. Collaboration w/ Yacine Jernite and David Sontag.

Data Science Intern, Comcast — Jun 2016 - Aug 2016
Analyzed large-scale user metrics to identify viewing behavior patterns.

Research Associate, Columbia Business School — Jul 2013 - Jul 2015

Selected Publications

Video Representation Learning with Joint-Embedding Predictive Architectures
K. Drozdov, R. Shwartz-Ziv, Y. LeCun. Preprint, 2024.
We develop a neural architecture that encodes object dynamics through self-supervised learning from video data. We incorporate variance regularization, which leads to improvements across multiple evaluation metrics.

Variance-Covariance Regularization Improves Representation Learning
J. Zhu, **K. Evtimova**, Y. Chen, R. Shwartz-Ziv, and Y. LeCun. Preprint, 2023.
We show that our regularization framework which encourages data representations to have high variance and low covariance enhances transfer learning in both the image and video domains.

Sparse Coding with Multi-layer Decoders using Variance Regularization
K. Evtimova, Y. LeCun. TMLR 2022.
ISTA is a classic algorithm for extracting sparse representations of data. We extend ISTA to work with deep neural networks, applying variance regularization to avoid collapse. Sparse image representations extracted with our method boost one-shot learning performance.

Emergent Communication in a Multi-Modal, Multi-Step Referential Game
K. Evtimova, A. Drozdov, D. Kiela, K. Cho. ICLR 2018.
We use reinforcement learning to train a multi-agent neural network architecture where agents cooperate to predict the class of an input image. The architecture is adaptive, using more computation for complex images.

Katrina Drozdov kve216@nyu.edu

Skills

LLM Post-Training: GRPO, reasoning diversity

Representation Learning: collapse-preventing regularization, sparse/EBM-inspired models

Deep Learning Stack: Python, PyTorch, Hugging Face ecosystem, experiment tracking with W&B

Training-at-Scale Foundations: distributed training basics

Selected Invited Talks

"Towards Building Intelligent Systems", Apple MLR — Oct 2024

"Deep Leaning", Leif Weatherby's course "Theory of the Digital" — Jan 2023

"Self-supervised Learning & Sparse Overcomplete Representations of Visual Data", CILVR at NYU — Jan 2020

Teaching & Mentorship

Mentor, New York University — Fall 2020
O. Che. Independent study on non-linear sparse coding.

Teaching Assistant, New York University — Spring 2020
Introduction to Machine Learning taught by Kyunghyun Cho at the Courant Institute.

Teaching Assistant, New York University — Spring 2019
Deep Learning taught by Yann LeCun at the Center for Data Science.

Teaching Assistant, Harvard College — Fall 2011
Linear Algebra and Applications taught by Vaibhav Gadre at the Math Department.

Professional Service

Conference Reviewing: ICML '21, '22, '23, '24; NeurIPS '21, '22, '25; ICLR '21, '22, '23, '24. AISTATS '24.

Additional Reviewing: TMLR (since 2024), WiML Workshop at NeurIPS '17.

Awards & Distinctions

Highlighted Reviewer, International Conference on Learning Representations (ICLR) — Apr 2022

Best Deep Learning Project Recipient, NYU Center for Data Science — Feb 2017

Ena Blyth Scholarship Recipient, Harvard College — Sep 2011 - May 2013

Leadership

Organizer, NYU AI School — Sep 2021 - Jan 2022

President, NYU Center for Data Science Leadership Circle — Sep 2018 - May 2019

Media Mentions

"From Academia to Industry: How a 2018 Paper Foreshadowed OpenAI's Latest Innovation" Medium, Oct 2024. Discusses my research on emergent communication and its connection to OpenAI's o1 model.

Interests

I enjoy singing and was a member of The Noteables, a Broadway show choir at Harvard. I also experiment with art and baking. My mixed media piece "Junk Mail" was featured at NYU's 4th Annual World Tour Pop-Up Gallery. Outside of my creative pursuits, I practice yoga and Tai Chi, and enjoy spending time outdoors, hiking in nature or walking around NYC.