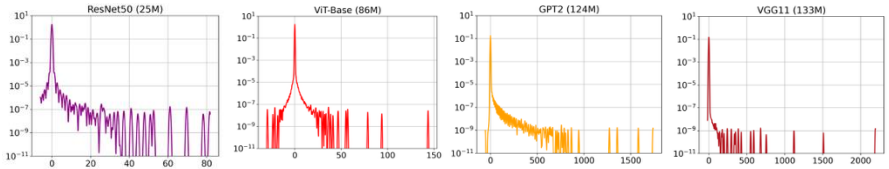
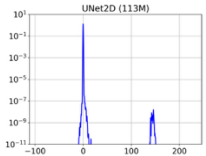
Deep Learning: “Low-Dimensional” Structures

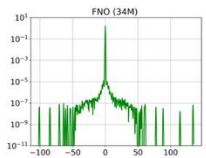
Loss Hessian eigen decay: Most eigen-values are at/near zero



Predictive Models: ResNets, ConvNets, Transformers (Vision, Language Models)



Generative Models:
Diffusion w/ U-Nets



PDE Solvers, Operators:
Fourier Neural Operator

Privacy in AI, Federated Learning, and Adaptation

Privacy in AI: Random Projection of Gradients + Gaussian Noise

- Sketched Gaussian Mechanism (SGM):
  - With clipped gradient $\tilde{g}_t$, do random projection, add noise

$$\hat{g}_t = R_t \tilde{g}_t + \xi_t, \quad \xi_t \sim \mathcal{N}(0, \sigma^2 \mathbb{I})$$

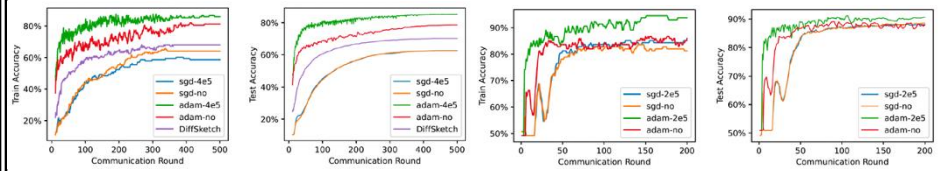
R is random projection
Projects high-d gradients to low-d

  - Update parameters, using OPT (SGD, Adam, AMSGrad, etc.)

$$\theta_{t+1} = \mathrm{OPT}(\theta_t, \hat{g}_t, \eta_t)$$

Noise is added to projected gradients

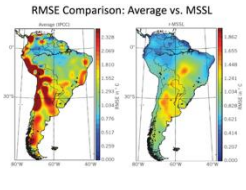
100x Random Projection (adam-xe5): Preserves Privacy, Outperforms in Accuracy



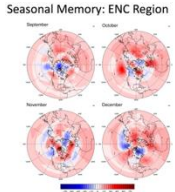
Federated SGM on Vision Models

Federated SGM on Language Models

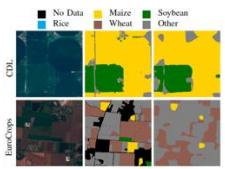
Applications: Weather, Agriculture, GeoSpatial AI



Weather (Multi-Task Learning)

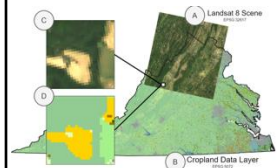


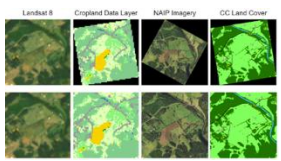
Seasonal (Interpretable AI)



Global Crop-Type (Foundation Model)

TorchGeo: GeoSpatial AI





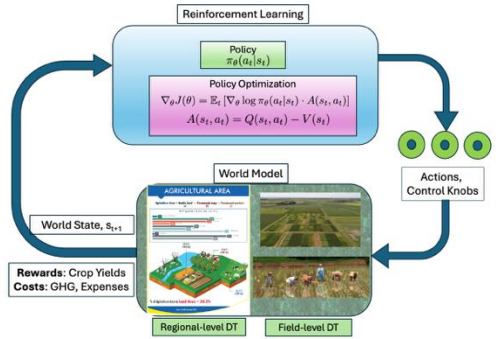
TorchGeo Impact (so far)

New geospatial domain library
- Easily load benchmark datasets
- Compose uncurated datasets
- Multispectral transforms
- First MSI pre-trained models

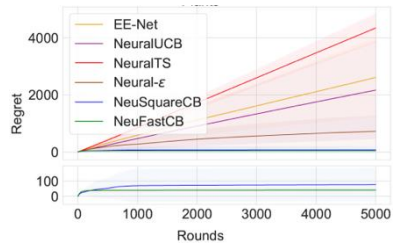
Widespread adoption
- 2.7K ⭐, 338 🍴, 70 contributors
- 10K PyPI downloads in past month (250K total)
- Researchers across academia and industry

EPFL · University of Cambridge · Stanford University · Carnegie Mellon University · Berkeley · MIT · University of Washington · ETH zürich · University of Oxford · intel. · IBM · Microsoft · Meta · aws

Sequential Decision Making under Uncertainty



Reinforcement Learning for Ag:
Field-level and Regional-level



Contextual Bandits:
Neural Policies
Exploration vs. Exploitation