Foundation Models

Technical Advances, Social Responsibility, Applications

Pls: Percy Liang, Russ Altman Jeannette Bohg, Akshay Chauhari, Chelsea Finn, Tatsunori Hashimoto, Daniel Ho, Fei-Fei Li, Chris Manning, Tengyu Ma, Chris Ré, Rob Reich, Dorsa Sadigh, Matei Zaharia

Applications

Household robotics dataset
- Goal: data from many distinct environments
- 70 people from 18 organizations
- 50+ real households with 20 scenes each
- 30K demonstrations (goal: 100K)

Voltron (robotic foundation model)
- Trained on Something-Something v2 (220K human video clips)
- Use text conditioning and generation to capture low-level spatial features and high-level semantics
- Evaluate on 5 diverse robotic tasks:
  - high-level semantics
  - capture low-level spatial features and

RoentGen (text-to-image model)
- Generates synthetic chest x-rays given natural language description
- Fine-tune Stable Diffusion on medical images
- Uses: data augmentation for improving classifiers, new teaching tool

Sophia: Second-order Clipped Stochastic Optimization
- Diagonal approximation of the Hessian (second-order) with clipping
- Outperforms Adam by 2x for LLMs from scratch on models from 125M to 6.7B parameters

LegalBench (evaluation)
- Legal reasoning: issue spotting, rule recall, analysis, conclusion
- 162 tasks (and growing!) contributed by ~40 legal experts across civil justice areas
- Evaluation on 20 FMs from 11 different families of models

Technical Advances

Domain Reweighting with Minimax Optimization (DoReMi)
- Uses distributionally robust optimization (DRO) to optimize domain weights on a small 280M proxy model
- Use domain weights to train a 8B model
- 2.6x faster (cheaper) than using heuristic domain weights

Hyena
- Long convolutions using Fast Fourier Transform
- O(n log n) instead of O(n^2)
- Result: 100x faster for long sequences

Sophia: Second-order Clipped Stochastic Optimization
- Diagonal approximation of the Hessian (second-order) with clipping
- Outperforms Adam by 2x for LLMs from scratch on models from 125M to 6.7B parameters

Social Responsibility

Holistic Evaluation of Language Models (HELM)
- A reproducible and transparent framework for evaluating foundation models.
- Leaderboards with many scenarios, metrics, and models

Hyena
- Long convolutions using Fast Fourier Transform
- O(n log n) instead of O(n^2)
- Result: 100x faster for long sequences

FM Ecosystem Graphs
- Foundation models encompass an ecosystem of datasets, models, and applications.
- This framework documents assets (datasets, models, applications) and their relationships within the foundation models ecosystem.
- Aids researchers, developers, policymakers, and the public in tracking foundation models, their creators, usage trends, and overall ecosystem dynamics.

Foundation Model Transparency Index (FMTI)
- Measures transparency of major developers on 100 indicators.
- Originally, developers scored an average of 37/100, but a follow-up showed improved scores: 14 developers averaged 58/100.
- Systemic opacity persists in areas like copyright status and data access.

Acknowledgments: We gratefully acknowledge the support of the Hoffman-Yee foundation for this work.