Depth First Learning: Learning to Understand Machine Learning

Avital Oliver
Google Brain
San Francisco, CA
avitalo@google.com

Surya Bhupatiraju
Google Brain
Mountain View, CA
sbhupatiraju@google.com

Cinjon Resnick
NYU
New York, New York
cinjon@nyu.edu

Kumar Krishna Agrawal
Google Brain
Mountain View, CA
agrawalk@google.com

Abstract

An experimental science is only as strong as the veracity of its experiments – when experiments are not reproducible, researchers struggle to place their work on solid footing and progress is hampered. Many fields have had reproducibility crises [5][8][3]. Some have recently recognized this as a problem in machine learning as well [12] [11].

One way to make research more reproducible is to make it easier to develop a deep understanding of the fundamentals underlying machine learning research papers. In this work, we present a new pedagogy, Depth First Learning, that addresses the challenges in understanding these fundamentals. In Depth First Learning, papers are studied down to their core conceptual dependencies, with additional material such as exercises and reproductions in code. We have used this approach in classes and codified our curricula as an open-source website available at `http://www.depthfirstlearning.com`.

1 Introduction

It can be hard to properly understand machine learning papers. Authors typically assume that you share the same context and prior knowledge that they have and write for a target audience of like-minded peers in their respective subfield. It is rare for papers to be written in a way that is accessible to an audience including both newcomers and experts. For newcomers, it is difficult to follow the motivations, results, and nuances of the paper. There may be caveats of the methods that are not clearly stated or unreported negative results that limit the generality of the results. Some try to understand a paper by additionally reading the cited papers. This magnifies the challenge as many of those papers assume yet a different set of context as the original paper.

The difficulties in understanding papers become especially apparent when one tries to reproduce the paper. Lacking a thorough understanding of the background knowledge can prevent the community from faithfully reproducing the work, especially when open-source implementations are not readily available [7]. Moreover, because there is often latent knowledge of how best to practically implement the algorithms that is not described in the research papers, it can be even more difficult to learn the methods from scratch. This affects the transition of research to application as well.

That being said, our goal here is not to argue that this is wrong or that the community needs to change the way that it conveys its work. On the contrary, we feel that this challenge is a consequence of

Preprint. Work in progress.

science. Rather, we introduce a pedagogical solution that we have developed over the past year to live alongside the traditional research paper ecosystem.

Depth First Learning is an approach to understanding modern research papers by taking a view that it is necessary to lay the foundations of understanding papers beyond only a superficial comprehension. We picked four different impactful works: InfoGAN [4], TRPO [14], AlphaGo Zero [15], and DeepStack [10]. For each of them, we carefully curated a curriculum that guides a student along a tree-like set of topic dependencies from the underlying concepts of the paper to the paper itself. We learned from reference material and the teachers who "wrote the book" where those resources were available. We wrote implementation code when it was applicable. We chose problems and we worked through them before meeting regularly to discuss. We learned about consequential papers not by one or two reads, but rather by working from the ground up to truly grasp the research direction and goals that the authors advanced. Along the way, we gained core competency in the topic, which we argue is also a key component to rapidly converting current research into both future research directions and industrial solutions.

2 Current Pedagogy

We want machine learning to be well understood and reproducible because then we are more confident in our footing to take the next big leaps.

One example of this issue emerging in machine learning is with the common Adam optimization algorithm [9] for stochastic gradient descent. A recent paper showed that the convergence proof in Adam, as originally constructed, was flawed, and demonstrated that the fix involves changing Adam itself[13]. This impactful paper was a stepping stone correction in science that had been missed by the entire field for more than four years, even though Adam had become the default learning optimizer in many practitioners' toolboxes.

How could we reduce occurences such as these in the future? We argue that increasing the understanding of machine learning fundamentals could help alleviate these issues. Towards this goal, one would understand what are the solid tree trunks of the knowledge base and what are the weak branches that should be tested. Our belief is that if there was an increase in focus on understanding fundamentals, then individuals in the field would have a stronger bent towards reliable and reproducible science.

As it stands, this is a challenging task. It is not only very difficult to drive yourself to learn in this way, but it is also a challenge to find a similarly keen community. What we see instead are the following existing pedagogy, and we compare them to our design decisions of Depth First Learning:

- **Textbooks** are often a de facto method of learning. We cannot and do not try to match the depth that they cover in a vast domain. Instead, Depth First Learning aims to to create a curriculum that provides a suitable depth for a single modern paper (while referencing chapters in textbooks).
- **Online Courses** provide a very different experience. While they are fantastic at reaching almost everyone in the world, they tend to organize around a similar, broad set of topics like textbooks do.
- **Blogs** are a copious resource in our community. They provide a counter to textbooks in that they frequently cover just a single paper or idea and can be a very welcome accompaniment to our own reading. In contrast to Depth First Learning, they generally are not a series of targeted learnings and are insufficient when it comes to deep understanding for anyone but the authors.
- **Distill** [1] publishes machine learning research with a focus on good, deep explanations, visualizations and interactive code environments. While there are similarities in our goals, Depth First Learning is not a place for publishing new research. Moreover, Depth First Learning does not have as much of an emphasis on interactive visualizations.
- **Metacademy** [2] is the closest parallel to Depth First Learning. We were inspired by this service and have similar goals in improving the learning process through breaking down concepts into their parents and gaining a more solid understanding. However, we ultimately have different foci as Depth First Learning is built around understanding significant machine learning papers and consequentially increasing the strength of the field itself.

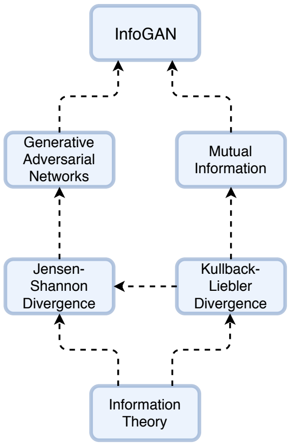

Figure 1: Dependency tree for InfoGAN[4]

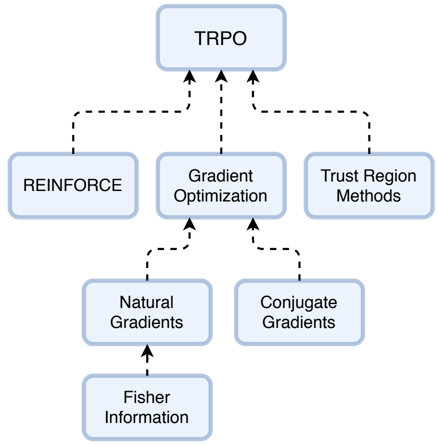

Figure 2: Dependency tree for TRPO[14]

3 Depth First Learning

What does a solution look like that addresses the challenges with the current approaches to learning, as described in the prior section? We present our solution, Depth First Learning.

One of most important criteria for understanding a paper is being able to position it within the context of other works. Otherwise, how could one tell if the work is incremental or is a great leap forward? How would a researcher know if they should reproduce this as a baseline? How would an industry player know if they should use this an early approach to solving their particular application? How would a researcher understand how to extend the work into new results? While papers usually contain exposition on previous work, it can still be difficult for the untrained eye to see past the caveats or shortcomings of the paper. It is often the case that the only reliable way to assess the novelty of the work is through reviews and conversations with experts in the field.

In light of these problems, we designed Depth First Learning to take a different approach from existing resources. We first build a dependency tree of concepts that influence the target paper and answer the following questions:

1. What minimal set of concepts do I need to understand the significance of this paper?
2. How do those concepts relate to each other and how do I progressively learn them?

When we implemented the ideas of Depth First Learning practically, we organized it as a recurring discussion group with a core set of organizers designing the curriculum and any interested participants. First, the core leaders for a given paper do a shallow read of that paper and produce an outline of the topics it covers. With this in hand, we trace these topics back to foundational concepts, like normal form games in game theory or policy gradients in reinforcement learning. They then find great material from which to learn those concepts, either by seeking general consensus or specific experts. Typically, the resources are notes, lectures, or book chapters, not more research papers. From that, they establish weekly discussion groups for a short course on the paper, where all the members are assigned to either reading or watching the relevant literature curated by the leads and are encouraged to come to each meeting with questions and discussion points. In some cases, the members are encouraged to answer questions, solve exercises, or implement specific algorithms or formulas from templates created by the leads.

We did the above for each of the four different papers mentioned: InfoGAN, TRPO, AlphaGo Zero, and DeepStack. We now describe these four papers in short, with the concepts they depend on, as were studied in our Depth First Learning sessions.

3.1 InfoGAN

InfoGAN[4] is an extension of GANs (Generative Adversarial Networks[6]) that learns to represent unlabeled data with meaningful codes, also known as representation learning. This is done by

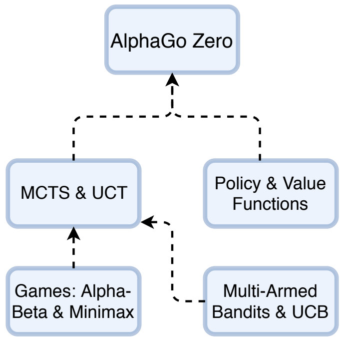

Figure 3: Dependency tree for AlphaGo Zero[15]

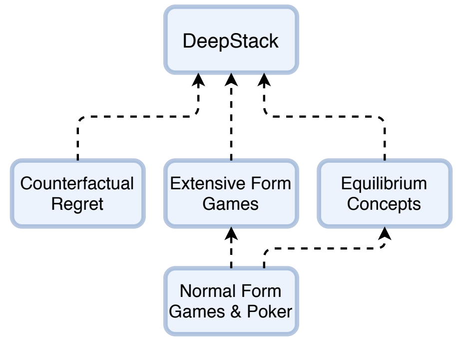

Figure 4: Dependency tree for DeepStack[10]

maximizing a lower bound on the mutual information between the generated samples and the codes. In order to understand the InfoGAN paper, one has to understand GANs and mutual information (a concept from information theory). Moreover, the GAN paper describes the optimization procedure as minimizing the Jensen-Shannon divergence (another information theoretical concept) between the data generating distribution and the distribution of samples from the network. Thus, information theory is used for both core dependencies of InfoGAN. See Figure 1 for a visual representation of this dependency graph.

3.2 TRPO

Policy gradient methods provide a model-free approach to training complex policies for reinforcement learning. Trust Region Policy Optimization (TRPO) [14] brings together insights from reinforcement learning and optimization theory to develop an algorithm which (under certain assumptions) provides guarantees for monotonic improvement and is often used as a strong baseline.

Understanding the TRPO algorithm includes first being familiar with a number of different subjects, including simple policy gradient methods, like REINFORCE, as well as different gradient optimization methods, like natural gradient descent, which in turn uses the conjugate gradient method. The latter additionally requires understanding the Fisher information matrix and how it is utilized. Finally, we study trust regions as a tool for optimization and understand how each of these concepts fit together to result in TRPO. See Figure 2 for a visual representation of this dependency graph.

3.3 AlphaGo Zero

AlphaGo Zero[15] was the first successful effort to definitively produce a machine that could beat humans at Go, while learning entirely from self-play. It is the culmination of many decades of understanding into games. The most important concepts used in the paper are Monte Carlo Tree Search (UCT algorithm) and policy and value functions from reinforcement learning. While the latter is fundamental to its field, for the former it is necessary to first understand multi-armed bandits, the UCB algorithm, and earlier game techniques such as alpha-beta pruning.

After developing an understanding of those techniques, absorbing the actual paper comes rather quickly as it relies on a variant of UCT to act as an expert that teaches a policy and value function through supervised learning. See Figure 3 for a visual representation of this dependency graph.

3.4 DeepStack

Where AlphaGo Zero developed an algorithm robust to solving many types of perfect information two player zero-sum games, DeepStack and its temporal sibling Libratus do the same for imperfect information two player zero-sum games.

In order to approach DeepStack, we start by understanding normal form games, like poker, which will lay the foundations for studying extensive form games and equilibrium concepts, along with their convergence proofs necessary to make DeepStack work. Finally, before understanding the main paper, we study counterfactual regret minimization (CFR), which is a recent algorithm developed to solve games of imperfect information. See Figure 4 for a visual representation of this dependency graph.

4 Discussion

We believe deeply understanding research ideas is essential for generating new research directions, understanding and critiquing current work, and being able to reproduce the work itself. Keenly aware of the material for a deep understanding into a modern paper, we created our own pedagogy for it - Depth First Learning. We kept careful track of all of our material at `http://www.depthfirstlearning.com`.

With Depth First Learning, we hope that the materials from textbooks, lectures, interactive code samples, online videos, discussion notes, etc. will help guide newcomers to a deeper understanding of the most important modern results in machine learning. We hope that it will serve as a platform for others to learn from and interact with the material.

Finally, we hope that others will join the effort by developing guides with us and contributing to our pool of papers to further the field of machine learning.

References

[1] Distill: http://distill.pub.

[2] Metacademy: http://metacademy.org.

[3] C. Glenn Begley and L. Ellis. Drug development: Raise standards for preclinical cancer research. *Nature*, 483:531–533, 2012.

[4] Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In *Advances in Neural Information Processing Systems*, pages 2172–2180, 2016.

[5] J. Leslie Glick. Scientific data audit—a key management tool. *Accountability in Research*, 2(3):153–168, 1992.

[6] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[7] Peter Henderson, Riashat Islam, Philip Bachman, Joelle Pineau, Doina Precup, and David Meger. Deep reinforcement learning that matters. *CoRR*, abs/1709.06560, 2017.

[8] Ioannidis JA. Contradicted and initially stronger effects in highly cited clinical research. *JAMA*, 294(2):218–228, 2005.

[9] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *CoRR*, abs/1412.6980, 2014.

[10] Matej Moravcík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael H. Bowling. Deepstack: Expert-level artificial intelligence in no-limit poker. *CoRR*, abs/1701.01724, 2017.

[11] Joelle Pineau. Reproducibility in deep reinforcement learning.

[12] Joelle Pineau. Reproducibility, reusability, and robustness in deep reinforcement learning.

[13] Sashank J. Reddi, Satyen Kale, and Sanjiv Kumar. On the convergence of adam and beyond. In *International Conference on Learning Representations*, 2018.

[14] John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In *International Conference on Machine Learning*, pages 1889–1897, 2015.

[15] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis. Mastering the game of go without human knowledge. *Nature*, 550:354–, October 2017.