THEORETICAL LIMITS OF DEEP LEARNING MODELS

Dr Krishna Murthy A
Assistant Professor
Department of Computer Science
Government First Grade College, Bhalki

Milindkumar M
Assistant Professor
Department of Computer Science
Government First Grade College, Bhalki

Abstract

An essential aspect of human motor learning is the formation of inverse models, which map desired actions to motor commands. Inverse models can be learned by adjusting parameters in neural circuits to minimize errors in the performance of motor tasks through gradient descent. However, the theory of gradient descent establishes limits on the learning speed. Specifically, the Eigen values of the Hessian of the error surface around a minimum determine the maximum speed of learning in a task. Here, we use this theoretical framework to analyze the speed of learning in different inverse model learning architectures in a set of isometric arm-reaching tasks.

It has often been argued that we face a trade-off between accuracy and opacity in deep learning models. The idea is that we can only harness the accuracy of deep learning models by simultaneously accepting that the grounds for the models' decision-making are epistemically opaque to us. In this paper, we ask the following question: what are the prospects of making deep learning models transparent without compromising on their accuracy? We argue that the answer to this question depends on which kind of opacity we have in mind. If we focus on the standard notion of opacity, which tracks the internal complexities of deep learning models, we argue that existing explainable AI techniques show us that the prospects look relatively good. But, as it has recently been argued in the literature, there is another notion of opacity that concerns factors external to the model. We argue that there are at least two types of external opacity—link opacity and structure opacity— and that existing techniques can to some extent help us reduce the former but not the latter.

Keywords:

Theoretical limits, Deep learning, Neural networks, Expressive power, Approximation theory, Universal approximation theorem, Sample complexity, Generalization bounds, Over fitting and under fitting, Bias–variance tradeoff, Computational complexity, Optimization landscape, Non-

convex optimization, Vanishing and exploding gradients, Data efficiency, Model capacity, Curse of dimensionality, Interpretability limitations.

Introduction

Humans are capable of producing fast and accurate movements to skillfully execute a variety of motor tasks. To accomplish such tasks, the central nervous system (CNS) uses feed forward mechanisms, as it cannot rely solely on delayed sensory feedback to guide the execution of movements.

Inverse models must be able to adapt to changes in the environment or the body to maintain successful task execution. This adaptation faces two main challenges: (1) errors in task space do not directly inform how motor commands should be adjusted to eliminate the errors, and (2) the models are generally not unique because of the redundancy of the motor system. Accordingly, several computational models such as direct inverse modeling, distal learning, and feedback error learning have been proposed as mechanisms to solve these challenges. These models differ in their architectures and mechanisms to adequately relate task errors to changes in motor commands. However, despite differences in their theoretical underpinnings, these models are all learned based on attempts to minimize an error quantity, such as performance error or motor error. Thus, from a computational perspective, the inverse learning problem can be formulated as a function fitting problem, where the strengths of neural connections in the inverse model circuit are the parameters that are tuned to fit the function.

There is evidence that deep learning models already by now outperform human experts in terms of accuracy and reliability in data heavy domains such as medicine. there are studies showing—at least in non-clinical settings—that deep learning models can perform just as well as experienced professionals when it comes to predicting breast and skin cancer based on diagnostic imaging.

Discussion

Based on the theoretical results of our computational framework, we simulated learning in four isometric arm tasks: (1) anisotropic scaling task, 2) anisotropic target distribution task, 3) virtual surgery task, and 4) target number task. In the anisotropic scaling and virtual surgery tasks we control the shape of the manipulability ellipse of the arm in different conditions. In the anisotropic target distribution and target number tasks, we control the target distributions in different conditions. The virtual surgery and target number tasks were included to explain experimental findings reported in previous studies in the context of our theoretical results.

For practitioners, issues of professional responsibility and legal liability are central. Practitioners have an obligation to make informed decisions and to fully understand the tools they utilize in patient care. Relying on an opaque model may hinder their ability to justify clinical decisions, explain outcomes to patients, and uphold professional standards. Moreover, legal accountability plays a significant role. If an adverse outcome arises from a decision based on an opaque algorithm, practitioners may face legal repercussions without the ability to demonstrate due diligence or articulate the reasoning behind their choices. This lack of transparency can challenge their role as informed decision-makers and erode confidence in the technologies they employ. For patients, trust and comprehension are critical factors influencing their acceptance of medical advice. Patients are more likely to be active participants in their healthcare decisions and feel assured about the safety and efficacy of proposed interventions when they understand the reasoning behind medical recommendations that impact them.

Explainable artificial intelligence To better understand what AI techniques can and cannot do, let us begin by introducing the idea of a black box deep learning model in some more detail. So consider a deep learning model that is designed to predict whether individuals are at risk of developing throat cancer. Suppose the model makes its predictions based on information from a huge and detailed medical dataset. More specifically, assume that each row in the dataset represents an input vector v for the model. Each row thus contains all the feature values for a particular data subject

Given that the number of input features in such a deep neural network can reach thousands, and the number of weights connecting the nodes can reach millions, if not billions, it is unsurprising that the correlations between inputs and outputs in these models are often described as opaque. Indeed, for all the numbers in a sufficiently complex deep neural network to make sense to us, we would need to be able to parse and comprehend extremely complex mathematical functions involving, potentially, billions of arguments and parameters. As cognitively and computationally limited creatures, we are of course unable to do any such thing. Moreover, unlike classical ‘rule-based’ AI systems, we cannot rely on a set of antecedently well-understood principles, rules, or laws to explain the statistical correlations upon which a deep neural network bases its output. While we may well understand the general mathematics behind activation functions or back propagation, this understanding does not extend to comprehending the billions of specific node values and connections within the network. Even if we could grasp all these details, it is unclear whether such knowledge would provide meaningful insight into why the model produces the outputs that it does

The dependence on Deep Learning models has been rising immensely for the past decade, due to their revolutionary performance on various problems in different domains. The performance gap between DL algorithms and older ML algorithms is very prominent in areas such as Image Classification, Object Detection, and Natural Language Processing.

We owe this great boost in performance to the differentiating features of Deep Learning:

1- The performance of DL models increases with the size of data.

2- DL models require a lot of computation power.

One of the main advantages regarding the performance of Deep Learning models is that they can be used to learn huge amounts of data. This makes up a distinctive aspect of Deep Learning, which sets it apart from other Machine Learning algorithms that we had been using before the emergence of Deep Neural Networks.

Literature review

Accordingly, the burden of calculating the weights and biases of a deep neural network is what proves to be the main reason the deep learning models are might be approaching their limits. It was known for a very long time that neural networks were computationally expensive. Yet, the progression of the hardware enabled us to utilize such power-consuming architectures. The progression of CPUs was good enough to keep our models to smaller-scale ones. Our DL models were able to be larger-scale with the coming of GPUs. The computations sped up to 35x, however, the growth speeds of architectures were even greater than the growth rate of computation power of GPUs. This may be one of the major problems in our present day, possibly setting a limit to our Deep Learning Models.

Early on it was clear that computational requirements limited what neural networks could achieve. In 1960, when Frank Rosenblatt wrote about a three-layer neural network, there were hopes that it had “gone a long way toward demonstrating the feasibility of a perception as a pattern recognizing device.” But as Rosenblatt recognized, “as the number of connections in the network increases, the burden on a conventional digital computer soon becomes excessive.” Despite this potential workaround, much of the academic work in this area was abandoned because there simply wasn’t enough computing power available. Since the growth in computing power per dollar closely mimicked the growth in computing power per chip, this meant that the economic cost of running such models was largely stable over time. Despite this large increase, deep learning models in 2009 remained “too slow for large-scale applications, forcing researchers to focus on smaller-scale models, or to use fewer training examples. But image recognition was just the first of these benchmarks to fall. Soon, deep learning systems also won at object detection, named-entity recognition, machine translation, question answering, and speech recognition. The introduction of GPU-based (and later ASIC-based) deep learning led to widespread adoption of these systems. But the amount of computing power used in cutting-edge systems grew even faster--at approximately 10× per year from 2012 to 2019. Instead, much of the increase came from a less economically attractive source: Running models for more time on more machines. It turns out that scaling deep learning computation by increasing hardware hours or number of chips is problematic because it implies that costs scale at roughly the same rate as increases in computing power, which will quickly make it unsustainable.

The introduction of GPU-based (and later ASIC-based) deep learning led to widespread adoption of these systems. But the amount of computing power used in cutting-edge systems grew even faster--at approximately 10× per year from 2012 to 2019. This rate far exceeded the ≈ 35 × total improvements gained from moving to GPUs-- meager improvements from the last vestiges of Moore’s Law – or the improvements in neural network training efficiency. Instead, much of the increase came from a less economically attractive source: Running models for more time on more machines. It turns out that scaling deep learning computation by increasing hardware hours or number of chips is problematic because it implies that costs scale at roughly the same rate as increases in computing power, which will quickly make it unsustainable.

Conclusion

In conclusion, we have developed a framework to theoretically identify factors that influence the speed of learning. This could potentially be used in more complex tasks to systematically identify body and environment parameters that facilitate learning, as well as to evaluate the difficulty of the task based on the elements of the task. Additionally, it could be used to systematically find candidate hypotheses about model architecture and components in the CNS to describe learning in a task.

The explosion in computing power used for deep learning models has set new benchmarks for computer performance on a wide range of tasks. However, deep learning's prodigious appetite for computing power imposes a limit on how far it can improve performance in its current form, particularly in an era when improvements in hardware performance are slowing. This paper shows that the computational limits of deep learning will soon be constraining for a range of applications, making the achievement of important benchmark milestones impossible if current trajectories hold. Finally, we have discussed the likely impact of these computational limits: Forcing deep learning toward less computationally intensive methods of improvement, and pushing machine learning toward techniques that are more computationally efficient than deep learning.

Reference

1. Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI).
2. IEEE Access, 6, 52138–52160. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., GilLópez, S., Molina, D., Benjamins, R., & Chatila, R. (2020).
3. Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. Information Fusion, 58, 82–115.
4. Bertone, G., & Tait, T. M. P. (2018). A new era in the search for dark matter. Nature, 562, 51–56. Bhansali,

5. A. & Jain, V. (2020) Using deep learning for healthcare resource allocation in COVID-19: Opportunities and challenges for LMICs. Center for Global Development.
6. J. C., & Busch, J. (2021). Artificial intelligence and patient-centered decision-making. Philosophy & Technology, 34, 349–371. Boge, F. J. (2022). Two dimensions of opacity and the deep learning predicament.
7. Minds & Machines, 32, 43–75. Buckner, C. (2018). Empiricism without magic: Transformational abstraction in deep convolutional neural networks. Synthese, 195(12), 5339–5372.
8. Buijsman, S. (2021). Acquiring mathematical concepts: The viability of hypothesis testing.
9. Mind & Language, 36(1), 48–61. Burrell, J. (2016). How the machine 'thinks': Understanding opacity in deep learning algorithms.
10. Big Data & Society, 3(1), 1–12. Carpenter, K. J. (1986). The History of Scurvy and Vitamin C. Cambridge: Cambridge University Press.
11. Char, D. S., Abràmoff, M. D., & Feudtner, C. (2020). Identifying ethical considerations for deep learning healthcare applications.
12. The American Journal of Bioethics, 20(11), 7–17. D'Amour, A., K. Heller, D. Moldovan,
13. C. Chen, J. Deaton, J. Eisenstein, M. Hoffman, et al. (2020).
14. Under specification presents challenges for credibility in modern machine learning. arXiv (arXiv: 2011.03395).
15. Explainable artificial intelligence (XAI) for increasing user trust in deep reinforcement learning driven autonomous systems.
16. A., Brunet, T. D. P., & Fisher, E. (2021). What is interpretability? Philosophy & Technology., 34, 833–862.