SAND2016-6341C

Overcoming the Static Learning Bottleneck - the Need for Adaptive Neural Learning

Craig M. Vineyard and Stephen J. Verzi
Sandia National Laboratories
Albuquerque, New Mexico 87185-1327
Email: cmviney@sandia.gov

***Abstract*—Amidst the rising impact of machine learning and the popularity of deep neural networks, learning theory is not a solved problem. With the emergence of neuromorphic computing as a means of addressing the von Neumann bottleneck, it is not simply a matter of employing existing algorithms on new hardware technology, but rather richer theory is needed to guide advances. In particular, there is a need for a richer understanding of the role of adaptivity in neural learning to provide a foundation upon which architectures and devices may be built. Modern machine learning algorithms lack adaptive learning, in that they are dominated by a costly training phase after which they no longer learn. The brain on the other hand is continuously learning and provides a basis for which new mathematical theories may be developed to greatly enrich the computational capabilities of learning systems. Game theory provides one alternative mathematical perspective analyzing strategic interactions and as such is well suited to learning theory.**

I. Introduction

A key differentiating capability of the brain is continuous learning. While humans may have courses and intentional moments in which we strive to learn content, fundamentally the brain is continuously re-wiring itself and making synaptic modifications as well as other forms of learning. Most machine learning algorithms however do not. Rather, most tend to have a fixed training phase and are then used as is. Machine learning algorithms do not have a lack of learning paradigms in fact there are many. For instance supervised and semi-supervised paradigms address how to handle labeled data. And methodologies such as batch, incremental, one-shot, and online address how data is presented to learning algorithms [1].

But despite this existent theoretical basis, learning theories have limitations, particularly with respect to dynamic environments. By design, the majority of machine learning algorithms are developed to converge upon a fixed solution. The optimization techniques fundamental to most learning algorithms (whether minimizing an error parameter or energy function) are designed to not only find a fixed state, but do so with provable convergence and stability. These desirable traits preclude the adaptivity of the algorithms. Instead of employing continual learning as brains do, many of these algorithms are more analogous to machine training. Not only do these limitations constrain algorithmic advances, but they furthermore limit architectural development in systems built without a theoretical basis. It is noted by Indiveri in [2] that, "For example, the models of neurons and synapses implemented on Neurogrid and BrainScales are hard-wired and cannot be changed" which in effect constrains neural inspired computing to be based upon hardware which in itself is also inflexible. Or as another example, compression techniques allow for vast reductions to the size of deep neural networks enabling their deployment on GPUs or neuromorphic chips rather than requiring a large computational cluster [3]. However, doing so comes at the expense of the adaptivity of the neural network as the pruned down result limits what the neural network may learn.

II. Learning Bottleneck

While clever techniques are being developed to address individual problems, a more holistic solution is needed. When game theory was pioneered by John von Neumann in the early 1940s [4], the motivating reason was that problems in economics were inadequately formulated with standard methods from optimization theory. Real world economic problems involving dynamic interactions were not adequately captured by single global objective functions and therefore needed a different approach [5]. An analogous statement can be said about machine learning — many learning problems involve dynamics not adequately captured by a traditional optimization approaches, but rather are limited by a 'Static Learning Bottleneck'. A static learning bottleneck in the sense that just as the memory access bottleneck limits the computational throughput of a von Neumann processor, likewise the static nature of machine learning algorithms limits what they may compute. Figure 1 illustrates this bottleneck in which distinct training and testing phases necessitate that for a model to be updated and learn it must be re-trained before it may be employed. Therefore to alleviate this bottleneck, a different adaptive learning approach is needed with alternative mathematics beyond optimization.

The brain however, employs continuous adaption and learning through a variety of forms of neural plasticity. Synaptic plasticity involves the dynamic alteration of the strength of the connections between neurons and includes mechanisms such as spike-timing dependent plasticity (STDP) [6] and short-term plasticity. Structural plasticity involves the addition and elimination of neural network infrastructure and includes mechanisms such as neurogenesis [7] and homeostatic plasticity [8]. These mechanisms have a variety of time scales enabling different properties such as the long term stability of the network as well as the short term encoding of novel

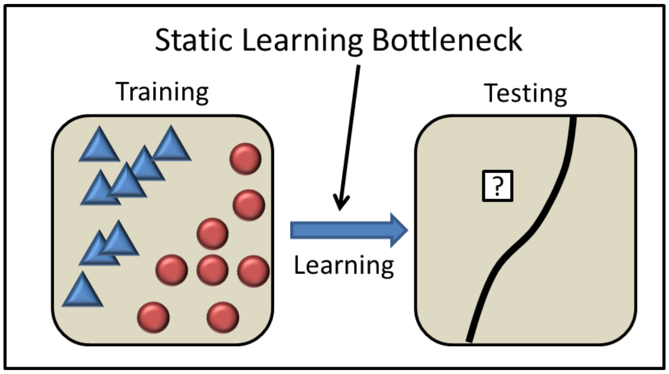

Fig. 1. Static Learning Bottleneck: the learning process as a limiting factor in the canonical training-testing paradigm of machine learning. By requiring a model be trained separately from being used, this makes the learning process a bottleneck

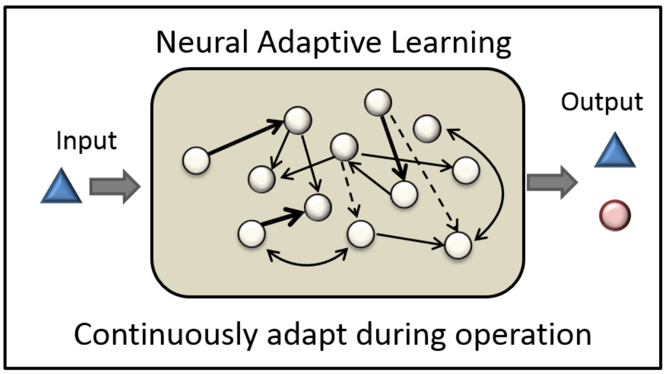

Fig. 2. Learning in the brain is a continuous and adaptive process rather than a disjoint paradigm with distinct training and testing periods. Rather learning occurs continuously throughout operation

stimuli. Effectively, neural plasticity mechanisms such as these and others amount to a spectrum of adaptive learning not captured by the operational mathematics or the learning paradigms typically employed in machine learning and neural network algorithms. But rather, they provide a foundation upon which adaptive learning theory may be based. Figure 2 is illustrative of a continuous learning paradigm brains employ in which they operationally adapt without a bottleneck such as forming new connections (dashed connectivity arrows), strengthening and weakening connections (illustrated by the width of connectivity arrows), or employing different activity patterns to name a few neural learning characteristics.

III. Game Theoretic Approach

One example alternative mathematical perspective is that of game theory. Just as game theory provided an alternative approach to optimization in economics, as a branch of applied mathematics to formally analyze strategic interactions it is well suited to formalize neural plasticity mechanisms as well. For example - voting, auction, and coalition mechanisms may describe synaptic and structural interactions implicit in learning. And concepts such as strategy profiles, equilibria, and repeated game play are applicable to temporal dynamics. Additionally, game theory encompasses many innate properties such as independence and an iterative and distributive nature which are well suited traits for adaptive scenarios.

An example in which a game theoretic approach has been successfully applied to adaptive machine learning problems is the SVM Game [9]. This game theoretic formulation of the Support Vector Machine (SVM) algorithm is an online iterative algorithm by which repeated interactions between data points leads to the identification of the salient ‘support vectors’ from which a discriminant may be formed. As an iterative game, the algorithm is able to address the inclusion of addition data by playing further game iterations and updating accordingly without having to re-compute the entire problem. Effectively the algorithm is able to address concept drift scenarios as a particular type of dynamic learning problem in which there are changes to the problem data over time such as a streaming domain or due to variation in underlying data distributions. For example, 3 illustrates the result of the SVM Game adaptation of a linear discriminant into a piecewise linear discriminant. The top half of the figure shows an original linearly separable problem for which the two classes later expand as shown in the lower half of the figure (for example this may be due to receiving additional sensor recordings over time) resulting in a later piecewise linear solution. This game theoretic formulation is based upon coalitional dynamics which may be related to ensembles of neural activity.

IV. Conclusion

In summary, the lack of adaptivity in most learning algorithms results in a static learning bottleneck. To overcome this limitation, neural plasticity mechanisms and dynamics could provide a basis for which new mathematical theories of adaptive learning may be developed. Game theory serves as one such mathematical basis with multifaceted applicability to learning theory. From a descriptive standpoint, it is capable of analyzing and providing explanations for the behaviours exhibited across a breadth of learning algorithms. Furthermore, it may also be used as the underlying decision-making mechanism governing the strategic decision making processes that comprise a learning problem. Effectively, game theory provides a strategic perspective as a means of interpreting, understand, and implementing learning problems.

Advances in the theory of learning have the potential to greatly enrich the computational capabilities of learning systems. In particular, a richer understanding of the role of adaptivity not only will help overcome the static learning bottleneck, but additionally provides a foundation upon which architectures and devices may be built.

Acknowledgment

This research was possible in part by LDRD program support from Sandia National Laboratories. Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department

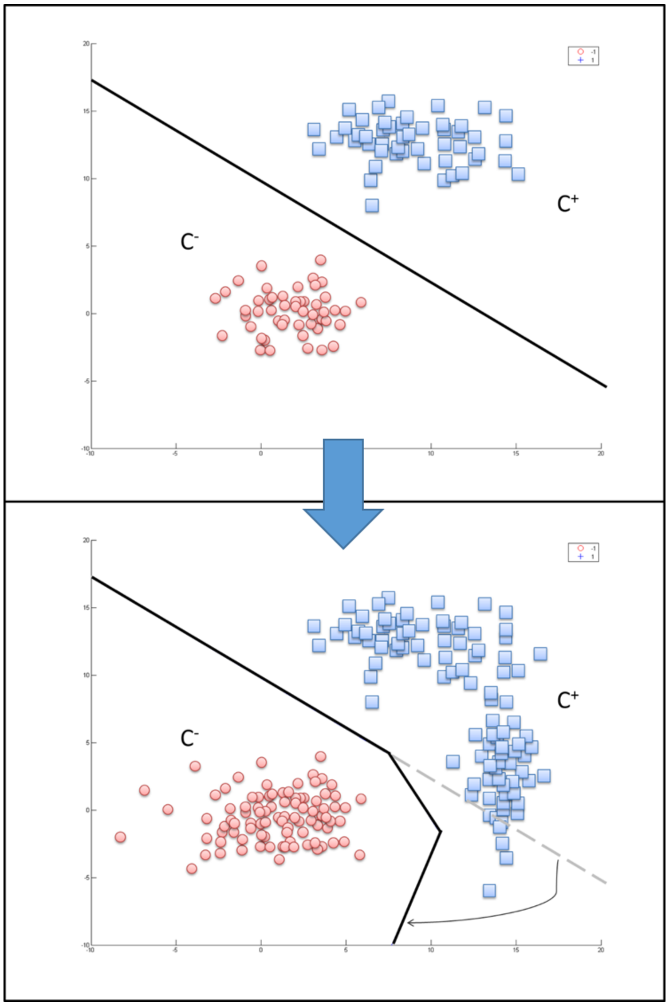

Fig. 3. An illustration of the SVM Game adapting a linear discriminant to a piecewise curved discriminant as a result of concept drift in the data

of Energys National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

[1] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of machine learning*. MIT press, 2012.

[2] G. Indiveri and S.-C. Liu, “Memory and information processing in neuromorphic systems,” *Proceedings of the IEEE*, vol. 103, no. 8, pp. 1379–1397, 2015.

[3] S. Han, H. Mao, and W. J. Dally, “Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding,” *arXiv preprint arXiv:1510.00149*, 2015.

[4] L. J. Neumann and O. Morgenstern, *Theory of games and economic behavior*. Princeton university press Princeton, NJ, 1947, vol. 60.

[5] M. Bicego, V. Murino, M. Pelillo, and A. Torsello, “Similarity-based pattern recognition,” *Pattern Recognition*, vol. 39, no. 10, pp. 1813–1814, 2006.

[6] G.-q. Bi and M.-m. Poo, “Synaptic modification by correlated activity: Hebb’s postulate revisited,” *Annual review of neuroscience*, vol. 24, no. 1, pp. 139–166, 2001.

[7] P. S. Eriksson, E. Perfilieva, T. Björk-Eriksson, A.-M. Alborn, C. Nordborg, D. A. Peterson, and F. H. Gage, “Neurogenesis in the adult human hippocampus,” *Nature medicine*, vol. 4, no. 11, pp. 1313–1317, 1998.

[8] G. G. Turrigiano and S. B. Nelson, “Homeostatic plasticity in the developing nervous system,” *Nature Reviews Neuroscience*, vol. 5, no. 2, pp. 97–107, 2004.

[9] C. M. Vineyard, S. J. Verzi, C. D. James, J. B. Aimone, and G. L. Heileman, “Repeated play of the svm game as a means of adaptive classification,” in *Neural Networks (IJCNN), 2015 International Joint Conference on*. IEEE, 2015, pp. 1–8.