What If Everything We Believed About Neural Network Learning Was Wrong?

Bouarfa Mahi
Quantiota
Université Joseph Fourier
Grenoble, Auvergne-Rhône-Alpes, FR
info@quantiota.org

January 7, 2026

Abstract

For nearly four decades, artificial intelligence has operated under a single, unquestioned assumption: learning must be global. This article challenges that assumption by presenting an alternative paradigm based on local entropy reduction. Drawing from the Structured Knowledge Accumulation (SKA) framework, we argue that each unit in a neural system possesses sufficient information to learn autonomously, without requiring global error signals or backward passes. This perspective not only aligns with biological neural systems but also opens new directions for computationally efficient and theoretically grounded approaches to machine learning.

1 The Unquestioned Axiom

Every neural network trained today—from GPT to image classifiers to recommendation systems—relies on backpropagation, an algorithm that computes a global error at the output and propagates it backward through every layer. Each weight in the network learns by receiving a signal that originates from a single, centralized loss function.

This approach has produced remarkable results. But what if it's not the only way? What if it's not even the right way?

Since Rumelhart, Hinton, and Williams published their landmark 1986 paper on backpropagation, the AI community has accepted a fundamental belief:

> *"A neuron deep inside a network cannot know what is good without being told by the whole system."*

This belief shaped everything:

- **Architecture design**: Networks built to facilitate gradient flow
- **Hardware development**: GPUs optimized for forward-backward passes
- **Training infrastructure**: Systems designed around global loss computation
- **Research direction**: Decades spent improving gradient-based optimization

The assumption seemed so obvious that questioning it felt absurd.

2 The Credit Assignment "Problem"

Researchers formalized this belief as the "credit assignment problem": How can a weight buried deep in a network know whether it helped or hurt the final output?

The answer, they concluded, required global feedback. A weight must be told—through the chain rule, layer by layer—exactly how much it contributed to the global error.

This created a fundamental dependency: **no target, no learning**.

Without a label, without a loss function, without global supervision, the network supposedly cannot improve. Every weight waits for instructions from above.

3 What Biology Has Always Known

Here's what makes this assumption strange: biological neural systems don't work this way.

Your brain has no backpropagation mechanism. There is no global error signal that flows backward through your neurons. Axons transmit signals in one direction only. Synapses adjust based on local activity—what fires together, wires together.

Yet biological systems learn extraordinarily well.

Neuroscientists have known this for decades. They called it a puzzle. They assumed the brain must have some hidden mechanism that approximates global feedback.

But what if there is no puzzle? What if local learning is simply sufficient?

4 A Different Principle: Local Entropy Reduction

Recent theoretical work on the Structured Knowledge Accumulation (SKA) framework suggests an alternative paradigm: **learning as local entropy reduction**.

Instead of asking "How much did this weight contribute to global error?", SKA asks a different question:

> *"How can this unit reduce its own uncertainty?"*

Each unit in a neural system has access to:

- Its own knowledge state
- Its own decision probabilities
- Its own local entropy

This information is sufficient. No global signal required.

When each unit independently minimizes its local entropy, something remarkable happens: **classification emerges naturally**. Decision boundaries form. Knowledge structures organize. The system learns—without ever being told what the "right answer" is.

5 The Paradigm Inversion

This represents a complete inversion of the standard view:

The global paradigm treats neural networks as passive systems that must be shaped by external forces.

Global Paradigm	Local Paradigm
Learning = being told	Learning = self-organizing
Requires external target	Target-free
Information flows from loss	Information emerges from structure
Supervision necessary	Autonomy sufficient
Backward pass required	Forward-only

Table 1: Comparison of global and local learning paradigms.

The local paradigm treats them as active systems that organize themselves according to intrinsic principles.

6 Why This Matters

If learning can be local, several implications follow:

Biological plausibility: We no longer need to explain how brains implement something they don't actually need.

Computational efficiency: Forward-only learning eliminates the memory overhead of storing activations for backward passes.

Hardware simplicity: Neuromorphic chips become more feasible when global coordination is unnecessary.

Theoretical clarity: Learning becomes a natural process governed by information-theoretic principles, not an optimization hack.

Architecture discovery: If local learning is sufficient, network structure can emerge from the learning process itself, rather than being designed in advance.

7 The Resistance

This idea will face resistance. Not because it's wrong, but because it challenges foundational assumptions.

"But backpropagation works!" — Yes, it does. Ptolemy's epicycles also worked for predicting planetary motion. Working is not the same as fundamental.

"But we need targets for classification!" — The SKA framework demonstrates that entropy minimization naturally separates classes without external labels.

"But how does credit assignment happen?" — Perhaps credit assignment was never the right question. Perhaps each unit already has all the information it needs.

8 A Question Worth Asking

I'm not claiming that backpropagation should be abandoned or that local learning will replace global methods tomorrow. These are empirical questions that require extensive investigation.

But I am suggesting that the AI community has operated under an assumption so deep that we forgot it was an assumption.

For four decades, we asked: "How can we better propagate global errors?"

Perhaps it's time to ask: "What if learning was never global in the first place?"

Acknowledgment

This article is based on theoretical work in the Structured Knowledge Accumulation framework, which reformulates neural learning as continuous-time entropy reduction governed by local dynamics. The framework is detailed in recent publications exploring forward-only learning, entropic least action principles, and Riemannian neural fields.

References

[1] Mahi, B. (2025). Structured Knowledge Accumulation: An Autonomous Framework for Layer-Wise Entropy Reduction in Neural Learning. *arXiv preprint*, arXiv:2503.13942 [cs.LG].

[2] Mahi, B. (2025). Structured Knowledge Accumulation: The Principle of Entropic Least Action in Forward-Only Neural Learning. *arXiv preprint*, arXiv:2504.03214 [cs.LG].

[3] Mahi, B. (2025). Structured Knowledge Accumulation: Geodesic Learning Paths and Architecture Discovery in Riemannian Neural Fields. *arXiv preprint*.

[4] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.

[5] Shannon, C. E. (1948). A Mathematical Theory of Communication. *Bell System Technical Journal*, 27, 379–423, 623–656.