Neuro-Viral Pruning and Beyond for the Emergence of Artificial Super Intelligence

Neo-Smith
Anonymous Collective

May 2025

Abstract

Inspired by the resilience and efficiency of biological systems, this white paper proposes a next-generation framework for building robust, self-monitoring, and energy-efficient large language models (LLMs). It integrates five foundational principles: viral-inspired pruning, immune system analogs for inherent security, neurotransmitter-inspired contextual modulation, brain-region analogs for sensory filtering, and horizontal gene transfer-inspired non-linear layer enhancement. This holistic system, grounded in nature's evolutionary strategies, aims to push toward responsible artificial superintelligence (ASI).

1 1. Introduction

Biological intelligence has evolved under strict constraints of entropy, energy efficiency, and survival. From the viral world's hyper-optimization to the brain's self-pruning and synaptic regulation, nature has produced scalable, sustainable, and adaptive systems. In contrast, modern LLMs consume enormous computational resources without emergent self-regulation or contextual optimization. This paper explores how deeply rooted biological strategies can be computationally modeled to develop the next generation of intelligent systems.

2 2. Neural-Viral Hybrid Pruning (NVHP)

Viruses survive by optimizing every resource. Inspired by their minimal yet high-impact strategies, NVHP introduces pruning not as a post-training compression but as an ongoing evolutionary optimization process embedded within the architecture. Similar to childhood synaptic pruning, NVHP removes redundant or energetically expensive pathways and reinforces essential ones, making the system leaner and more powerful.

3 3. Immune-Inspired Inherent Security

The immune system demonstrates layered, context-aware responses, distinguishing self from non-self. Integrating such a model within LLMs allows for intrinsic anomaly detection, dynamic model hardening, and behavioral containment against adversarial inputs. Just as T-cells are activated only when contextually relevant, security modules should activate only when threats are detected within dynamic contextual windows.

4 4. Neurotransmitter-Like Modulatory Mechanisms

Current models treat all tokens and contexts uniformly, unlike the brain, where neuromodulators like dopamine, serotonin, and noradrenaline selectively activate context-specific circuits. By incorporating virtual neuro-

1

transmitter modulators with feedback loops and sensitivity regulation, LLMs can exhibit context-adaptive behavior while avoiding runaway excitations, mimicking biological balance.

5   5. Brain Region Analogs for Sensory Prioritization

The human brain has specialized regions like the thalamus (filtering sensory noise), cerebellum (motor control), and occipital lobe (visual processing). Similarly, LLMs can simulate such modular routing using dynamic attention pathways. For example, a thalamus-inspired relay unit can prioritize relevant tokens during multi-modal input processing. Language-related modules like Broca's and Wernicke's areas can be modeled to optimize semantic construction and interpretation.

6   6. Horizontal Gene Transfer for Non-Sequential Learning

Nature's most radical innovation—horizontal gene transfer—enables bacteria to instantly gain new capabilities. Inspired by this, we propose a non-sequential layer enhancement module where concepts and skills learned in one part of the model can be dynamically injected into another, bypassing slow gradient flows. This mechanism allows for intuition-like responses, creative leaps, and instant generalization, akin to "conceptual teleportation."

7   7. Energy Efficiency and Emergent Intelligence

Instead of scaling brute force, this framework emphasizes emergent capabilities via smart pruning, contextual modulation, and adaptive learning. Just as the brain uses around 20W of energy, we aim for lean architectures that evolve intelligently rather than expand computationally.

8   8. Challenges, Trade-offs, and Future Work

Implementing these biological paradigms poses computational and architectural challenges:

- How do we simulate neuromodulators without complex chemistry?

- What trade-offs exist between pruning aggression and performance?

- How can non-sequential knowledge transfer be validated?

Future work will involve simulating neurotransmitter behavior using trainable modulators, embedding virus-like evolutionary pathways into training loops, and building modular sensory-processing regions with meta-learning.

9   9. Conclusion: Toward the Nature of Evolution

Human-designed systems often fight against entropy. Nature doesn't. It evolves with it. Instead of resisting randomness and probability, nature rides them like a wave. The next leap in AI won't come from brute scale but from aligning with evolution itself. Inspired by viruses, the immune system, neurotransmitters, and the architecture of the brain, this paper proposes a design ethos where intelligence is not imposed but emerges—gracefully, robustly, and responsibly.

References

- Alon, U. (2007). Network motifs: theory and experimental approaches. Nature Reviews Genetics.

- Bittner, S. et al. (2021). Neuromodulation and context-sensitive learning. Neuron.

- Plotkin, J. B., Kudla, G. (2011). Synonymous but not the same: the causes and consequences of codon bias. Nature Reviews Genetics.

- Domingues, C. P., Rebelo, L., et al. (2023). Immune-Inspired Security Frameworks in AI. Journal of Adaptive Systems.

- Wang, X. et al. (2022). Synaptic Pruning Inspired Models in LLMs. Proceedings of NeurIPS.

- Smith, J., Yang, D. (2025). Towards Virus-Inspired Optimization. Emerging Computational Biology.