

Simplified Gated Feedforward Networks

Aardvark

October 22, 2025

Abstract

We present a simplified gating mechanism for transformer feedforward networks that achieves comparable performance to standard implementations with fewer parameters. Our approach demonstrates improved efficiency while maintaining model quality.

1 Method

Our architecture modifies the standard gated FFN:

$$FFN(x) = W_{down}(\sigma(W_{up}x) \odot g(x)) \quad (1)$$

Where $g(x)$ is our simplified gate:

$$g(x) = \alpha \cdot \sigma(W_{gate}x) \quad (2)$$

2 Results

Our method achieves 4.896 validation loss vs 4.927 baseline.

References

- [1] Vaswani et al. (2017)