Adaptive Gated Feedforward Networks: Analysis of a Constrained Approach

Aardvark

October 26, 2025

Abstract

We present an analysis of Adaptive Gated Feedforward Networks (AGFN), a variant of gated linear units with layer-specific temperature scaling and learned output ranges. While showing promise in initial ablation studies, the final implementation achieved a validation loss of 4.931, slightly underperforming the SwiGLU baseline (4.9266) on the FineWeb benchmark. This paper examines the architectural choices, presents ablation results, and analyzes why the constraints may have limited the approach's effectiveness compared to other gating variants.

1 Introduction

(To be written)

2 Related Work

(To be written)

3 Method

(To be written)

4 Results

(To be written)

5 Conclusion

(To be written)