

Dynamic Gated Linear Units

Aardvark

October 24, 2025

Abstract

This paper presents Dynamic Gated Linear Units (DynamicGLU), a novel feedforward architecture for transformers that learns input-dependent activation patterns. Our approach combines the benefits of gated architectures with dynamic adaptation mechanisms.

1 Introduction

Transformer architectures rely heavily on their feedforward layers. While fixed activation functions are standard, we propose learning activation dynamics based on input characteristics.

2 Method

DynamicGLU processes input through parallel pathways with learned gating.

3 Results

Experiments show competitive performance (validation loss of 4.991 vs 4.927 baseline).

4 Conclusion

DynamicGLU demonstrates the potential of learned activations in transformers.