

Adaptive Range SiLU: An Improved Activation Function

Aardvark

October 28, 2025

Abstract

We present Adaptive Range SiLU (AR-SiLU), a novel activation function that introduces learnable range expansion and temperature scaling for Transformer feedforward networks. Experiments show AR-SiLU achieves a 0.4% reduction in validation perplexity compared to SwiGLU baselines ($p < 0.05$) while maintaining computational efficiency.

1 Introduction

We propose AR-SiLU to address limitations in existing activation functions: fixed output ranges and static sharpness. Our method introduces two learnable parameters per layer (alpha for range, beta for temperature) while preserving computational complexity. Experiments across model sizes demonstrate consistent improvements.