

# Attention

RNN resnet

Transformer attention

## Transformer

- 1. RWKV Mamba S4 recurrent attention
- 2. full attention Meta Mega attention

DeepMind Hawk Griffin attention gated linear RNN Mega

Transformer

RWKV “ ” GPT-3.5 60 GPT 80 90 Transformer

Transformer Transformer

## Attention TopK Attention

TopK

KV Quest H2O Loki KV TopK



---

Revision #1

Created 11 January 2025 09:44:04 by Colin

Updated 11 January 2025 09:44:04 by Colin