■■■■■■■■■■■■

■■: Attention Is All You Need ■■: Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, ■ukasz Kaiser, Illia Polosukhin■Google Brain & Google Research■■■ ■■■■■■: 2024■

■■: ■■■■■■■■■

■■■■■■ChatGPT■■■■Midjourney■■■■■■AI■■■■■■■■■■■■■■■■■■■■■■■■■"■■■"——Transformer■

2017■■

■■

■■■■: ■■■■■■■■■

2017■■AI■■■■■■■■

■Transformer■■■■■■■■■■■■■■■■■——■■■■■■■■■■■■■■——■■■■■■■■**■■■■■■■RNN■**■■■■■■■■■**■■■■■■■■LSTM■**■

■■■■■■■RNN■■■■■■■■■■■RNN■■■■■■■■■■■■■■■■■■■■"■■"■■■■■■■"■■■■■■■"■■■"■■■■■"■■■■■■■"■■■■■■"■■■"■■■■"

■■■■■■

■■■■■■■■■■

■■RNN■■■——■■■■■■■■

[illegible]GPU[illegible]

[illegible]

[illegible]RNN[illegible]

[illegible]LSTM[illegible]"[illegible]"——[illegible]

[illegible]

[illegible]"[illegible]"[illegible]

[illegible]——"[illegible]"[illegible]"[illegible]"[illegible]

■■■■■■■■**■■■■■■Attention■**■

■■■■■■■: ■■■■■

■■■■■■■

■■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■■■■■RNN■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■"■■■"■■■■■■■■■■■■■■■■■■■■■■

■■■■■■"■■■"■■■■■■■■■■■■■■■■■■■■■■■■

■Transformer■■■■■■■■■■■■■■■■**■■■■■■■■**■■■■■■■■■■■■■■■■■■■■■■■■■■■■■"■■■■■■■■■■■■■■■■■■■■■■"

■■■"■■■"■■■■■■**■■■■■■■■■■■■■■■■■■■■**■

Transformer■■■■■

■■■■■Transformer■■■■■■■■■■**■■■■Encoder■**■**■■■■Decoder■**■

■■■■■"■■"■■■■■■■■■■■■■■■■"■■"■■■"■■■"■

■■■■■■■■■RNN■■Transformer■■■■■■**■■■■**■■■■■■■■■■■■■■■■■■■■■■"Attention Is All You Need"■■■■

■■■■■■■■■■■

■■■■■■■■■■■■**■■■■■■■Self-Attention■**■

■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■■■■"■■■■■■■■■■■■■■■■■■■■■"

■■■■■■"■"■■■■■■■■■■■■"■"■■■■■■■■■■■■■"■"■"■■■"■■■■■■■■■■■■■■"■■"■"■"■■■■

■■■■■■■■■■■■■■■■■■■■■ - **■■■Query■**■■■■■■■ - **■■Key■**■■■■■■ - **■■Value■**■■■■■■■■■■■

"■"■■■■■Query■■■■■■Key■■■■■■■■■■■■■■■■■"■■"■Key■■■■■■■■■■■■"■■"■Value■

■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■RNN"■■"■■■■

■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■■"■"■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■**■■■■■■Multi-Head Attention■**■■■■■■■■■■■■■■■■■■8■■■16■■■■■■■■■■■■■■■■8■■■■■■■■■■■■■■■■■■■■■■■■■■

■■

■■■■■■■■■■■■■■■■■■

■■■

■■■■■■■■■■**■■■■■Positional Encoding■**■■■■■■■■■■■■■"■■■■"■■■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■1■2■3■4■■■"■■■■"■■■■■■■■■■■■■

[illegible]

[illegible]"[illegible]"[illegible]"[illegible]"[illegible]

Transformer[illegible]**[illegible]**[illegible] `output = x + f(x)`[illegible]

[illegible]"[illegible]"[illegible]

[illegible]**[illegible]**[illegible]

[illegible]: [illegible]

[illegible]

■■■■■■■■■Transformer■■■28.4■BLEU■■■■■■■■■■■■■■■■■■■■■■■■■■2■■■■

2■■0.1■■■■■■■■

■■■■■■■■■■■■■■■■41.8■■■■■■■■■■■■■■

■■■■■■■■■■■■

■■■■■■■■■■■■■■■■■

■■■

■Transformer■■■■■■■■8■GPU■■12■■■■■■■■■■3.5■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■AI■■■■■■■■■■

■■■■■■■■■■■■■■

■■■RNN■"■■■"■■Transformer■■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■"making...more difficult"■■■■■■■■■■■■■■■■"making"■"difficult"■■■■■■■■■■■■■■Transformer■■■■■■■■"■■"■■■■■■■

■■■■■■■■■■■■■■■

■■■■Transformer■■■"■■■■"■■■■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■"■■■■■■■"■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■Transformer■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■Transformer■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■■■■■ - ■■■■■■■■■■■■■■■"■"■■■■■■■ - ■■■■■■■■■■■■■■■■■■■■ - ■■■■■■■■■■■

■■

> ## ■■■■: ■■■■■■

■■■■■

■Transformer■■■■■■■■■■■■■■■■■■RNN■■■■■■■■■■"■■■■RNN"■■■LSTM■GRU■■■RNN■■■■■■■RNN......

Transformer■■■■■■■■■RNN■

■■

■■■■■■

■■■■■■■■■2010■■■■■■■■■■■■■GPU■■■■GPU■■■■■■■■RNN■■■■■■■■■■■■■GPU■■■■

Transformer■■■■■■"■■■■■■"■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■GPT-3■■■■■GPU■■■■■■■■■Transformer■■■■■■■■■■■■■■■

■■■■■■

Transformer■■■■■■■■■■■■■■"■■"■

RNN■■■■■■■■■■■■■■■■■■■■■■■■■■■Transformer■■■■■■——■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■"■■■■■■■"■■■■■■GPT-3■1750■■■■GPT-4■■■■■■■■■■■■■■■■Transformer■

■■■■■■■

■Transformer■■■■■■■■RNN■■■■■■CNN■■■■■■■GNN......■■■■■■■■■■■■■■■■■

■Transformer■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■Vision Transformer(ViT)■Transformer■■■Transformer■■■■

■■■■■

■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■■■■■■■$O(n^2)$■■■n■■■■■■

■■■■■■■■■■■■1000■■■■■■■■■1000×1000=100■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■■■■100■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■

■■■■■■■■■■■■■■■■■■■■■■■ - **■■■■■**■■■■■■■■■■■■■■■■■■"■■■"■■ - **■■■■■**■■■■■■■$O(n^2)$■■O(n) - **■■■■**■■■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■■■Claude■GPT-4■■■■■■■■■■■■■■■■■■■■

■■■■■■■

■■■■■■■■■■Transformer■■■■■■■■■■■

1. **■■■■■■■■**■■■■■■■Transformer■■■■■■■GPU■■■■■■■■■■■■■■■■

2. **■■■■■■**■■■■■■■■■■■■■■■■■■■■■"■■"■■■■■■■■■■

3. **■■■■■■**■Transformer■■■■■■■■■■■■■■■■■■■■■■■■

■■■■

■■■■■■■■■■■■■■■■■■■■■■■■■■■■■"■■■■"■■■■

■2017■■■■"■■■■■■■RNN"■■■AI■■■■■■■■■■■■■■■■■■RNN■■■■■■■■■■■■■■■■■■■■■■■■■■■■RNN■■

■■■■■■■■

■■■

■■■■■■■■■■■■■■■■■■■Transformer■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■"■■"■■■■■■■■■AI■■■■■■

■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■

■■■■■■■■■■■■■■■■■■■■■■■■■■■■■■AI■■■■■■■AI■■■■■■■■■■■■■■■■■■■■■■■■2017■■■■"■■"■■■

■■■■■■■■■■■■■■■■■■■■■"All You Need"■■■■■■■■■

■■

2017■■■■■■■■■■■■■Transformer■■

■■■ ■ ■■■■: ■■■■/■■■■■■/■■■■ ■■ ■■■■: ■15■■ ■■ ■■■■: ■■■■■■■■■■■■■■■■

■■■ ■ ■■■■: `papers/downloaded_paper.pdf` ■■ ■■■■: 102.8■ ■■ ■■■■: ■■■API ■■■ ■ ■■■■: claude-opus-4-5-20251101 (via Claude Agent SDK) ■ ■■■■: 2026■01■17■ 11:33:57

■■■■ GitHub Actions + Claude Agent SDK + ■■■■ ■■■■