

Transformers

- It emerged in 2017, the article 'Attention is All You Need' by Vaswani et al.



Transformers

Revolutionary Model

Used in GPT-3/4, T5, and BERT

Doesn't use recurrence like RNN

Based on the attention mechanism

Encoder/Decoder architecture

Attention / Self Attention

- Attention learns the contextual relationship between words.

