Learning Science - Sample

Spacing, attention, and how memory works.

7 flashcards · Knowledge Compounder study guide

Transformers introduced self-attention in 2017 with the paper 'Attention is All You Need'. The key insight is that you c

1. What is self-attention in the context of Transformers?

Self-attention is a mechanism that computes each output position as a weighted combination of all input positions, allowing the model to directly represent long-range dependencies in a sequence without relying on recurrence.

2. How does scaled dot-product attention work to produce position weights?

Each input position is projected into query, key, and value vectors; the dot product of a query with all keys (scaled to prevent large magnitudes) is passed through a softmax to produce attention weights, which then weight-sum the value vectors to produce the output for that position.

3. When would you choose a Transformer architecture over a recurrent network?

Reach for a Transformer when you need to model long-range dependencies in sequences and require massively parallel training across many hardware accelerators, since its attention mechanism processes all positions simultaneously rather than sequentially.

Spaced repetition works because of the spacing effect: memories consolidate during the gaps between reviews, not during

4. What is the spacing effect in spaced repetition?

The spacing effect refers to the phenomenon where memories consolidate during the gaps between reviews, not during the reviews themselves.

5. How does successful recall affect the forgetting curve?

Successful recall just before forgetting occurs resets the forgetting curve to a shallower slope.

6. When should you review material to optimize long-term retention?

You should review material less often but at the right moments, as determined by algorithms like SM-2 and FSRS that estimate memory stability and desired retention.

7. What are SM-2 and FSRS in spaced repetition?

SM-2 and FSRS are algorithms used in spaced repetition to schedule the next review of a card based on its estimated memory stability and the user's desired retention.