

Transformer Feed-Forward Layers Are Key-Value Memories

[link](#)

Рецензия (Денисенко Наталья)

Вклад статьи: исследование менее изученного модуля трансформеров - feed-forward layers.

Сильный стороны:

- интересная идея
- множество экспериментов
- сравнение с паттернами собранными людьми

Слабые стороны: отсутствуют

Стиль написания:

- порядок повествования понятен
- все исследуемые модули разъяснены
- формулы и эксперименты описаны понятно
- графики довольно аккуратные

Оценка: 9 и 4