

Vanilla self attention

Linearized self attention