

Layer 0 Attention Weight Distribution

0.50

Attention Weight Value

Layer 1 Attention Weight Distribution

0.25

Sticky Tokens Normal Tokens

0.75

Sticky Tokens Normal Tokens

1.00

0.25

0.20

0.15

0.10

0.05

0.25

0.20

0.00

Frequency (%)