

Heatmap: Ambiguous and CoLA layers

COCO

mean

mean

mean

EMMT

cls

sum

mean

CoLA

mean

sum

sum

BERT

GPT2

SBERT

12

10

8

6

4

2