Heatmaps of attention logits $\frac{1}{\sqrt{d'}}\langle Qx_i,Kx_j\rangle$ imdb_bert-large-uncased_params-random_dense-on_layers-24

