

bert-xlarge-v2_uniform-tokens_{x,y}_trained-params_dense-off_24-layers)

cos similarity heatmap - layer 16

