

cos similarity heatmap - layer 13
albert-xlarge-v2_wikitext_{Qx,Ky}_trained-params_dense-off_24-layers)

