

cos similarity heatmap - layer 3
albert-xlarge-v2_wikitext_{Qx,Ky}_trained-params_dense-on_24-layers)

