

cos similarity heatmap - layer 23  
albert-xlarge-v2\_wikitext\_{Qx,Ky}\_trained-params\_dense-off\_24-layers)

