

bert-large-uncased_wikitext_{Qx,Ky}_random-params_dense-off_24-layers)

cos similarity heatmap - layer 5

