

bert-large-uncased_wikitext_(x,y)_random-params_dense-off_24-layers)

cos similarity heatmap - layer 3

