cos similarity histogram - layer 13 (albert-xlarge-v2_wikitext_(Qx,Ky)_trained-params_dense-on_24-layers) 8 7 6 5 4 3 2 1

0.00

-1.00

-0.75

-0.50

-0.25

0.25

0.50

0.75

1.00