

bert-large-uncased_uniform-tokens_{Qx,Ky}_trained-params_dense-off_24-layer

cos similarity histogram - layer 7

