

Layer:model.layers.0.attention.wk.weight alpha=2.41, ks\_distance=0.04

