

Layer:model.layers.1.attention.wk.weight alpha=10.88, ks\_distance=0.12

