

Layer:model.layers.1.attention.wk.weight alpha=2.03, ks\_distance=0.08

