

Layer:model.layers.1.attention.wk.weight alpha=1.98, ks\_distance=0.05

