

Layer:model.layers.1.attention.wk.weight alpha=1.96, ks_distance=0.05

