

Layer:model.layers.1.attention.wk.weight alpha=8.44, ks_distance=0.15

