

Layer:model.layers.1.attention.wk.weight alpha=8.75, ks_distance=0.12

