

Layer:model.layers.0.attention.wk.weight alpha=4.72, ks_distance=0.21

