

Layer:model.layers.0.attention.wk.weight alpha=3.20, ks_distance=0.12

