

Layer:model.layers.0.attention.wk.weight alpha=2.56, ks_distance=0.04

