

Layer:model.layers.1.attention.wk.weight alpha=3.39, ks_distance=0.05

