

Layer:model.layers.1.attention.wk.weight alpha=2.51, ks_distance=0.04

