

Layer:model.layers.0.attention.wv.weight alpha=3.42, ks_distance=0.08

