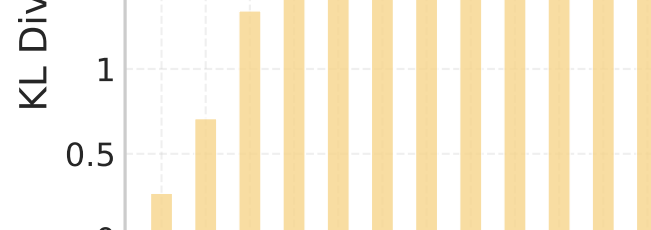


Wasserstein Distance Between Sticky and Normal Tokens Across Layers



KL Divergence Between Sticky and Normal Tokens Across Layers

