

KL Divergence Between Base and Fine-tuned Models During Training

