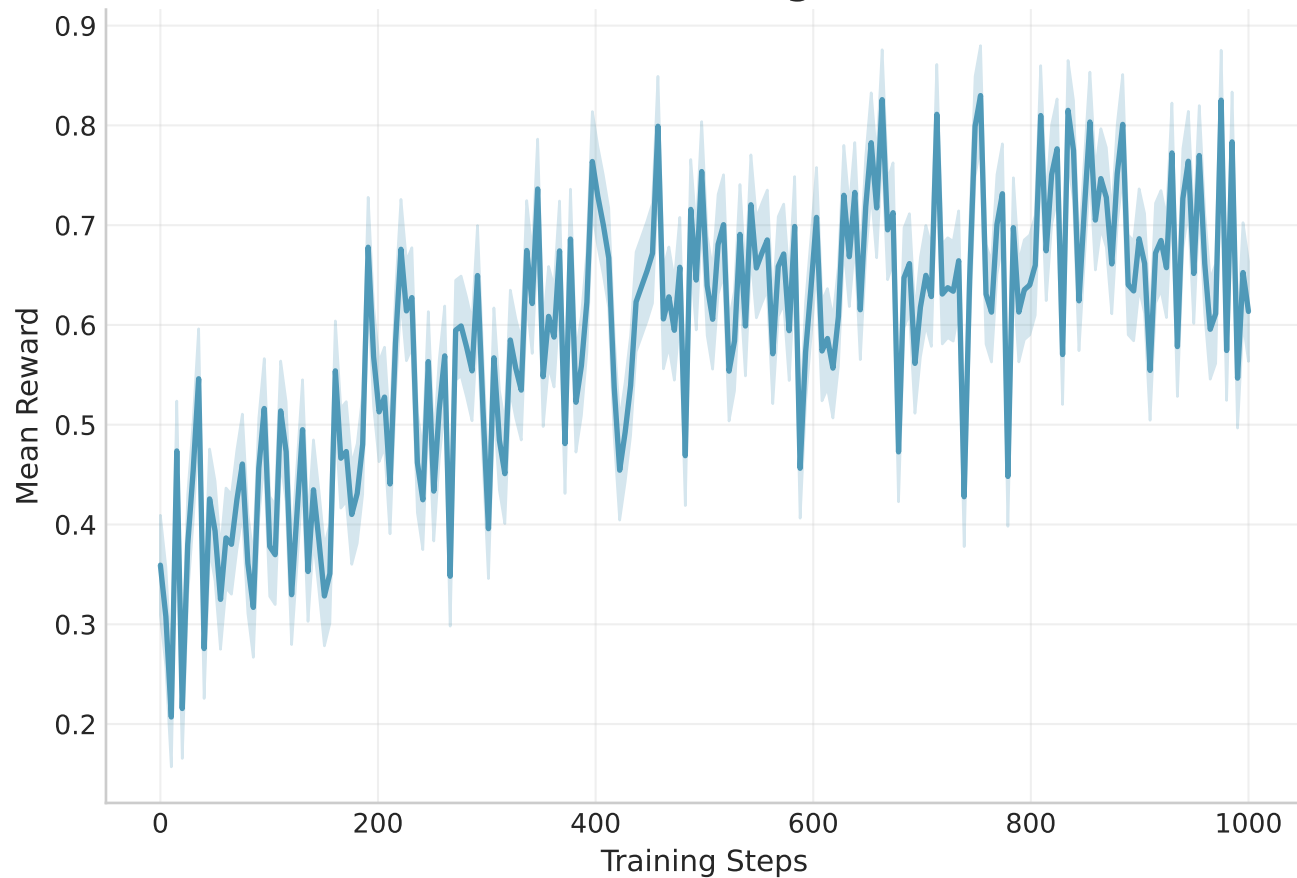
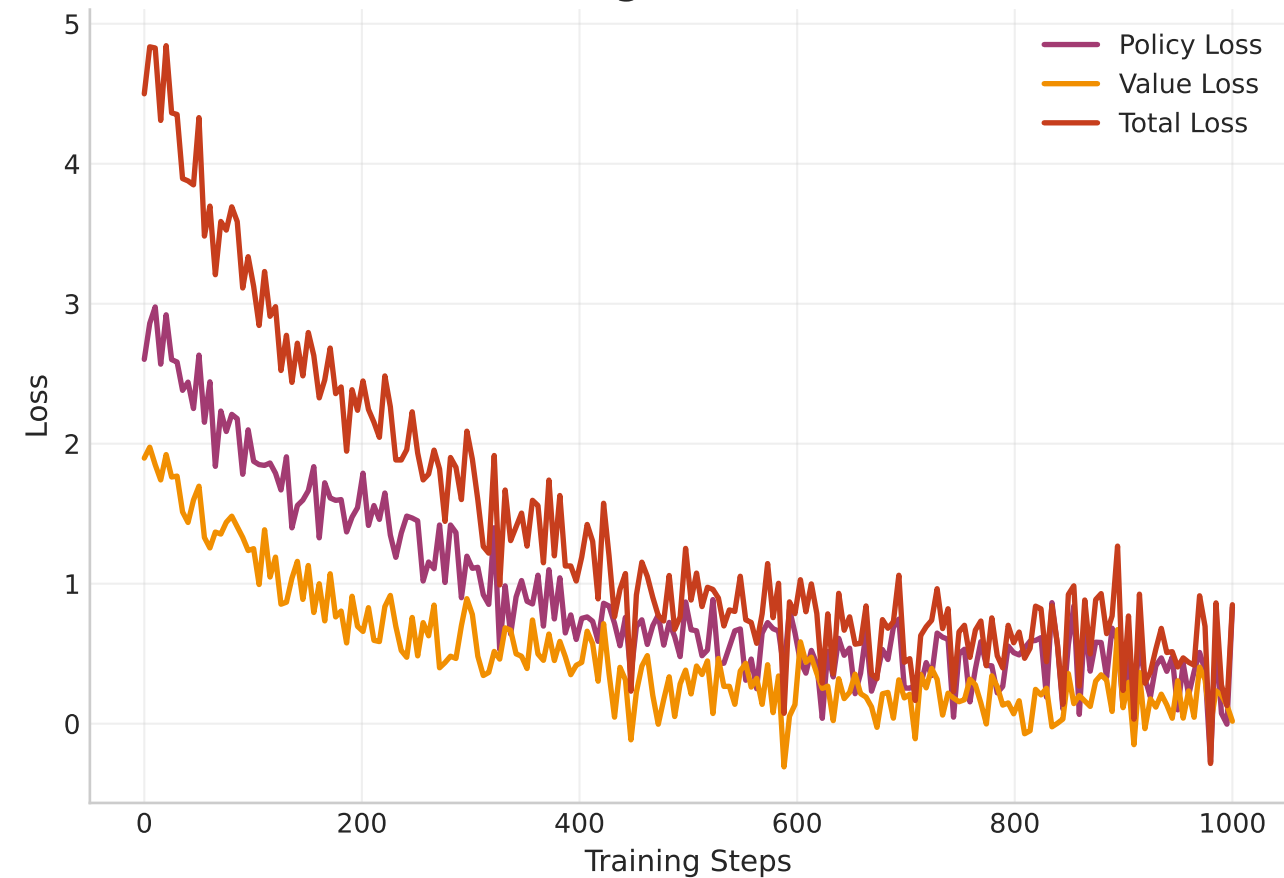


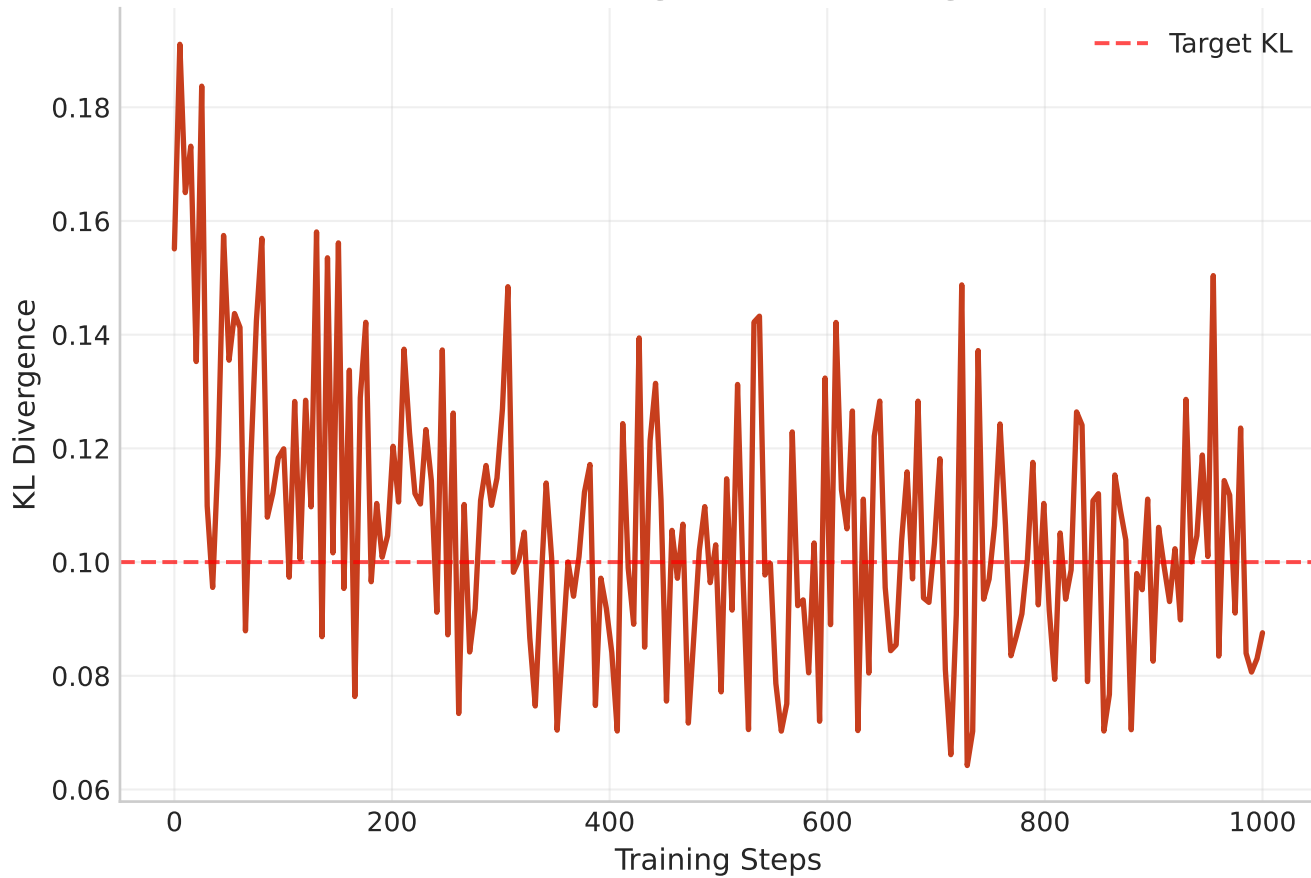
PPO Reward Progression



Training Loss Curves



KL Divergence Tracking



Policy Entropy Evolution

