

# PPO in Medical Applications

## Proximal Policy Optimization (PPO)

PPO is the most widely used algorithm for RLHF, providing stable updates that prevent catastrophic policy changes in medical AI systems.

### Key Features



#### Clipped Objective

Limits policy updates to prevent drastic changes



#### Trust Region

Maintains proximity to previous policy



#### Stability

Reduces training instability vs vanilla PG



#### Sample Efficiency

Reuses data multiple times per iteration

### Training Process (Iterative Loop)

1

2

3