Pusher

- Ours (finetuned)
- Ours (not-finetuned)
- Human Preferences (finetuned)
- Human Preferences (not-finetuned)
- PPO (finetuned)
- PPO (not-finetuned)