

Agent performance trajectories across datasets

