

Agent performance under distributional shift without test-time adaptation

