AMD Project Proposal. Using Imitation Learning to improve performance for the open source models in the BrowserGym leaderboard.

Doug Chang
dougchang25@gmail.com

Jonathan Oakey
xxx@gmail.com

Abstract

With the introduction of Claude Work on 1/12/26 it is being clear LLMs can be used for more than coding assistants. The multi agent frameworks being introduced are closer to substitutes for roles currently occupied by human employees. BrowserGym `https://huggingface.co/spaces/ServiceNow/browsergym-leaderboard` shows the effectiveness of LLMs in controlling browsers as agents. We would like to explore adding an open source LLM to the leaderboard using Imitation learning to increase performance vs. adding LLMs without fine tuning for this specific task.

1 Introduction

The specific problem is the lack of opensource models as an alternative to reducing cost and increasing performance for user specific tasks and domains. Most agents are calls to a commercial agent service. The BrowserGym paper appendix (Le Sellier de Chezelles et al. (2025)) lists costs at approximately $\$1k$ per $280 - 500M/3 - 40M$ input/output tokens. The top entry in the BrowserGym leaderboard is GenericAgent-Claude-3.7-Sonnet. This is a call to Claude-3.7. This doesn't offer any alternatives for performance and cost customization.

2 Methodology

We propose to break this task into several parts. The first part is to collect data which can be used in the supervised stage of behavorial cloning. Commercial browser traces are estimated to be in the 10^7 to 10^9 range.

The second stage is to run traces on customized distributed infrastructure. This may change depending on how data collection goes.

1. Use a nextJS app as a replacement for sandbox because chatGPT generated can quickly generate correct code and the deploy cycles are rapid. Kubernetes provisioning is slow and it takes more time to extract out metrics and make changes to the application. This has no value as open source it is generated by chatGPT and a productivity tool.

2. Synthetic Data: Behavior cloning is the naive IL strategy; it suffers from possible distribution shifts between training and test time. DAgger helps to compensate for the distribution shift. Commercial agents supplement training with additional synthesis strategies which we intend to explore and document

3 Benefits to AMD

There are several benefits to AMD as a infrastructure sponsor.

1. The possibility of increased cloud GPU business with AMD GPUs as solutions in multi agent solutions. Currently API calls to Claude/OpenAI dominate the agent market. Imitation learning should improve the performance of lower performing open source LLMs.

2. A version of vllm deployment software different than the current production stack. The current vllm deploy stack is a lightweight NextJS applicaiton starting and destroying droplets. This is closer to a lightweight sandbox than elastic kubernetes clusters. The current demo works on AMD CPU Droplets.

3. A reference solution for AMD on Imitation Learning and the various strategies of how to improve performance ranging from DAgger to trajectory synthesis.

References

Thibault Le Sellier de Chezelles, Maxime Gasse, Alexandre Lacoste, Massimo Caccia, Alexandre Drouin, Léo Boisvert, Megh Thakkar, Tom Marty, Rim Assouel, Sahar Omidi Shayegan, Lawrence Keunho Jang, Xing Han Lù, Ori Yoran, Dehan Kong, Frank F. Xu, Siva Reddy, Graham Neubig, Quentin Cappart, Russ Salakhutdinov, and Nicolas Chapados. 2025. The BrowserGym Ecosystem for Web Agent Research. *Transactions on Machine Learning Research* (2025). `https://openreview.net/forum?id=5298fKGmv3` Expert Certification.