

---

# **AMD Project Proposal. Using Imitation Learning to improve performance for the smaller 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. The cost and performance browser agents is a key part of the adoption of these frameworks

The specific problem is cost and performance of the smaller models on the browsergym leaderboard. Most agents are calls to a commercial agent service. The BrowserGym paper 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. There is no RL applied to this agent. We intend to use Imitation Learning to try to produce an equivalent faster and lower cost agent on this leaderboard.

## **1 Introduction**

The introduction of Claude work has quickly spawned an open source version, eigent. This was a real time pivot by a startup after the Claude announcement. This accelerates the experimentation landscape from BrowserGym leaderboard to a real multi agent framework.

## **2 Methodology**

We propose to break this task into several parts. The first part is to use IL to improve agent performance and register it on the BrowerGym leaderboard. This part includes using IL to improve leaderboard performance. If there is sufficient progress on the BrowerGym leaderboard we will expand to include multitask agent platforms such as eigent.

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.