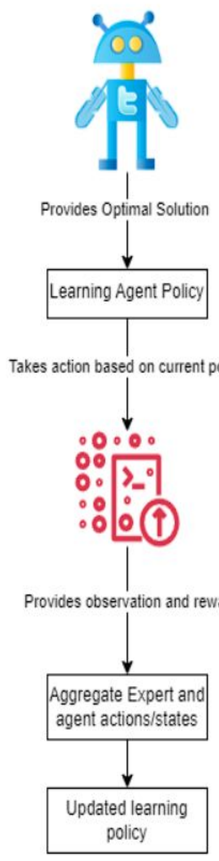
Leveraging Human Knowledge in Imitation Learning

Professor Srijata Das, Assistant Researcher Ibrahim Alwishah
College of Engineering & Computer Science, University of Michigan-Dearborn

SURE EXP+ Undergraduate Research

INTRODUCTION

- The research used an interactive learning framework to train an imitation learning agent with expert knowledge, enabling policy training with less data through iterative reinforcement learning.
- Our research demonstrated how DAgger improves AI decision-making and performance.



DAgger graphic

PURPOSE STATEMENT

- Purpose: Explore and implement DAgger and HG-DAgger algorithms in reinforcement learning.
- Framework: Use OpenAI Gym.
- Goal: Address limitations of traditional imitation learning by improving policy performance with DAgger in unfamiliar scenarios.

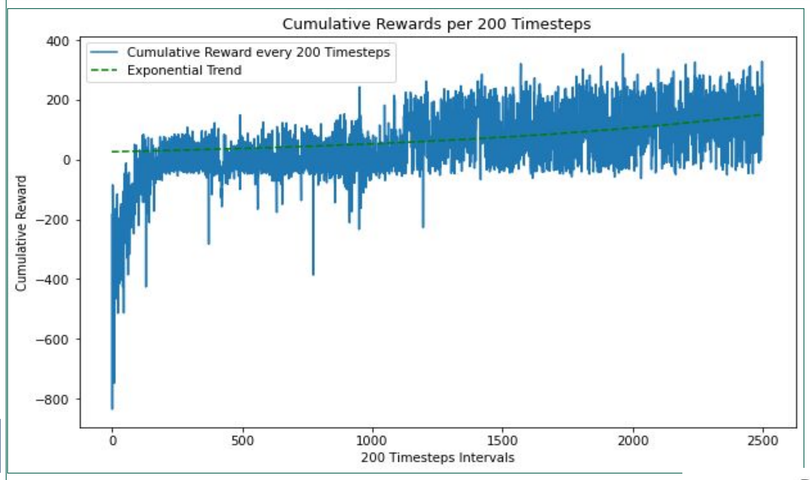
METHODS

- Trained a pseudo human using DQN to act as expert in the environment with the purpose of guiding a less experienced AI agent.
  - Integrated the expert into the DAgger framework to iteratively guide the AI agent.
- Trained the agent with DAgger to refine decision-making
  - Used a benchmark library to iteratively improve the policy with expert feedback.
- Evaluated the data by tracking average rewards
  - Data is used to demonstrate performance improvement as the agent learned from expert feedback over time.

DAgger Algorithm [1]

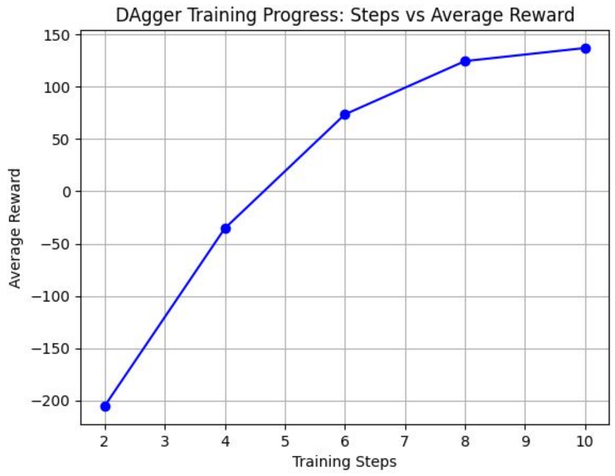
Initialize $\mathcal{D} \leftarrow \emptyset$.
Initialize $\hat{\pi}_1$ to any policy in Π.
for $i = 1$ **to** N **do**
  Let $\pi_i = \beta_i \pi^* + (1 - \beta_i)\hat{\pi}_i$.
  Sample T-step trajectories using π_i.
  Get dataset $\mathcal{D}_i = \{(s, \pi^*(s))\}$ of visited states by π_i and actions given by expert.
  Aggregate datasets: $\mathcal{D} \leftarrow \mathcal{D} \bigcup \mathcal{D}_i$.
  Train classifier $\hat{\pi}_{i+1}$ on $\mathcal{D}$.
end for
Return best $\hat{\pi}_i$ on validation.

RESULTS



On the left is the process of training the Pseudo human using DQN

On the right you can see how DAgger leverages expert guidence and feedback to ease the training process



SKILL : CRICITAL THINKING

- Evaluation: Assessing the validity of various sources of information and methods.
- Model Training: Used Stable-Baselines3 and other libraries for developing and evaluating reinforcement and imitation learning models.
- Data Analysis: Tracking and evaluating performance metrics, such as average rewards, to assess improvemen

CONCLUSIONS

- The purpose of DAgger: Understood that these algorithms are designed to improve traditional imitation learning.
- Importance of Iterative learning: Gained a greater understanding of how DAgger collects new data during each iteration.

Future Work

- Uncertainty Estimates: Experiment with various types of uncertainty estimates to enhance model robustness.
- Explainability: Work on making DAgger more explainable to better understand decision-making processes.
- Model Variations: Test different models, including transformers and Graph Neural Networks, to explore their effectiveness.

Reference

[1] “A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning.” Ross et al. 2011