Reinforcement Learning for Self-Driving Cars (Seasons of Code 2025)

**Name:** *Ganesh Sonawane*

**Objective:** To train an autonomous driving agent using reinforcement learning on the highway-v0 environment using PPO, evaluate its performance, and study the effects of modifying environment parameters such as traffic density and lane count.

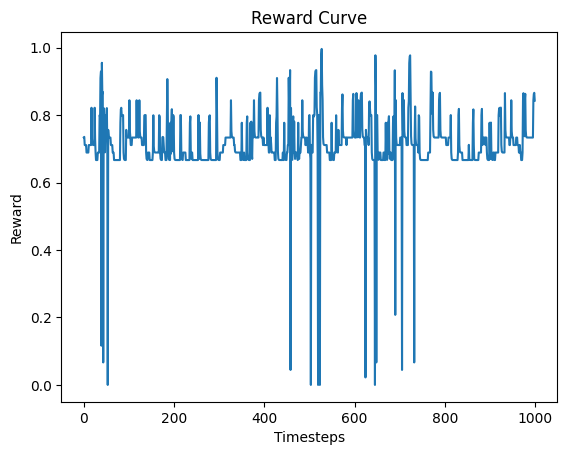
**Setup and Algorithm:**

* **RL Algorithm: PPO with MlpPolicy**
* **Environment: highway-v0 (gymnasium)**
* **Training Library: Stable-Baselines3**
* **Device Used: CPU (Colab)**
* **Training Steps: 10,000 timesteps**

**Metrics:**

|  |  |
| --- | --- |
| **Metric** | **Value** |
| **Total Timesteps** | **10,000** |
| **Total Training Time** | **77.27 Minutes** |
| **FPS** | **2** |
| **Mean Episode Reward** | **20.7** |
| **Mean Episode Length** | **28.7** |
| **Crashes (in evaluation)** | **0** |

**Reward Curve:**

****

**Simulation Video:**

****

**Insights Challenges:**

* Training was significantly slower on CPU (2 FPS), taking ~ 77 minutes for 10k steps.
* The agent began to avoid collisions and drive more safely after ~5k steps.
* Rendering during training slows it down — disabled it for better speed.
* Future improvements could include:
* Custom reward shaping (e.g., penalize lane switching)
* Curriculum learning: gradually increasing traffic
* Trying SAC or DDPG for continuous control

**Github Repository Link:** [**https://github.com/Cosmox999/SOC-RL**](https://github.com/Cosmox999/SOC-RL)