AI Lunar Lander

Alex Smith and Shreya Modi
CS 4100

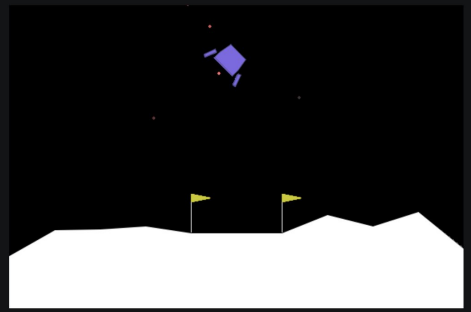
GOAL

Train a model to successfully land the lander in the target area.

What is our environment?

Gymnasium

An API standard for reinforcement learning with a diverse collection of reference environments



A

Actions

- Left Thruster
- Right Thruster
- Main Thruster (up)
- Nothing

B

State

Horizontal Position (x)

Vertical Position (y)

Horizontal Velocity (vx)

Vertical Velocity (vy)

Angle (θ)

Angular Velocity (ω)

Left Leg Contact (l)

Right Leg Contact (r)

AGENT

Deep Q Network

$$Q(s,a) = Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a') - Q(s,a))$$

HYPERPARAMETERS

Learning rate α , Gamma γ, Epsilon ε,

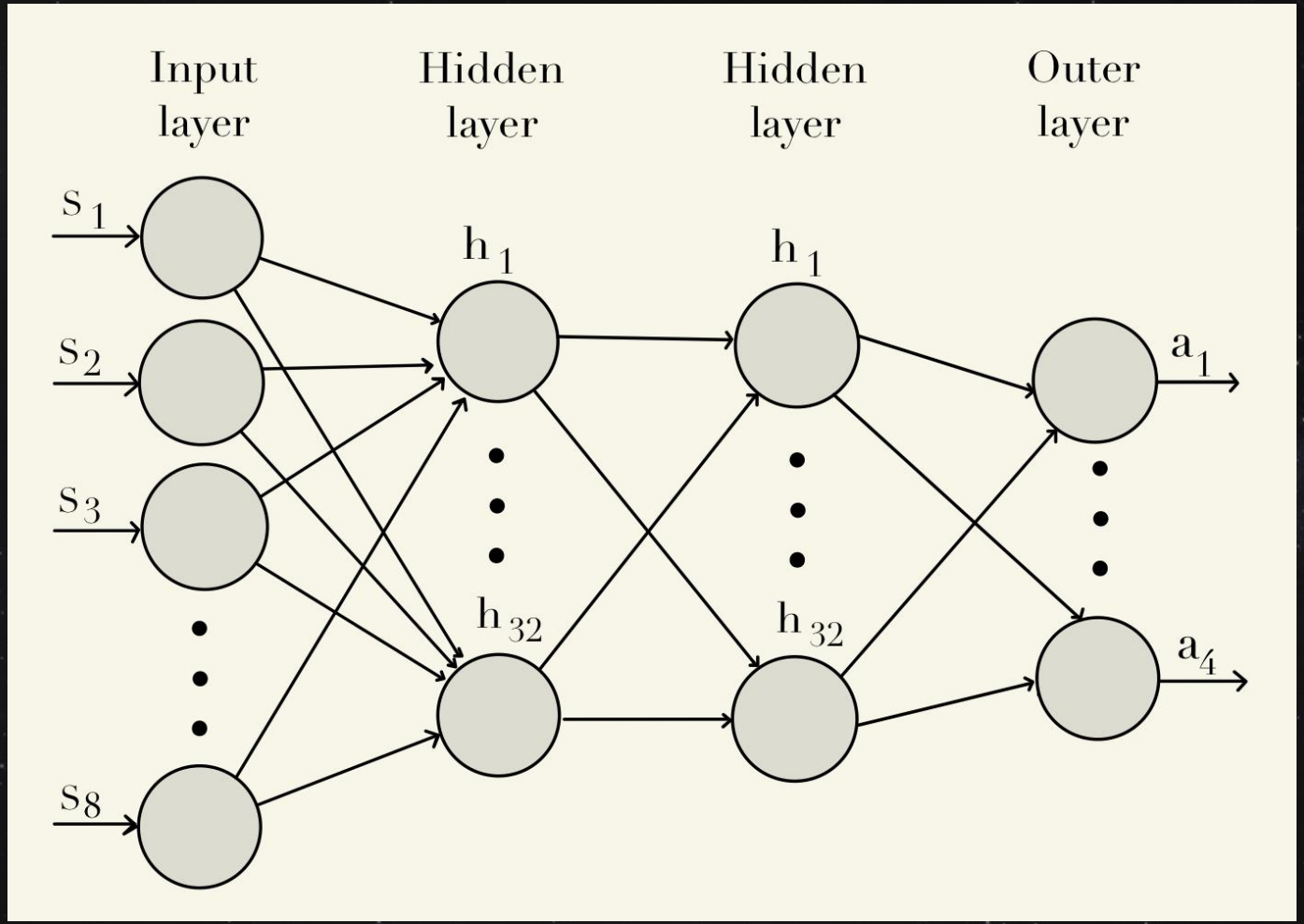
Epsilon Decay Rate,

Update_Frequency,

Hidden layer width,

Replay buffer

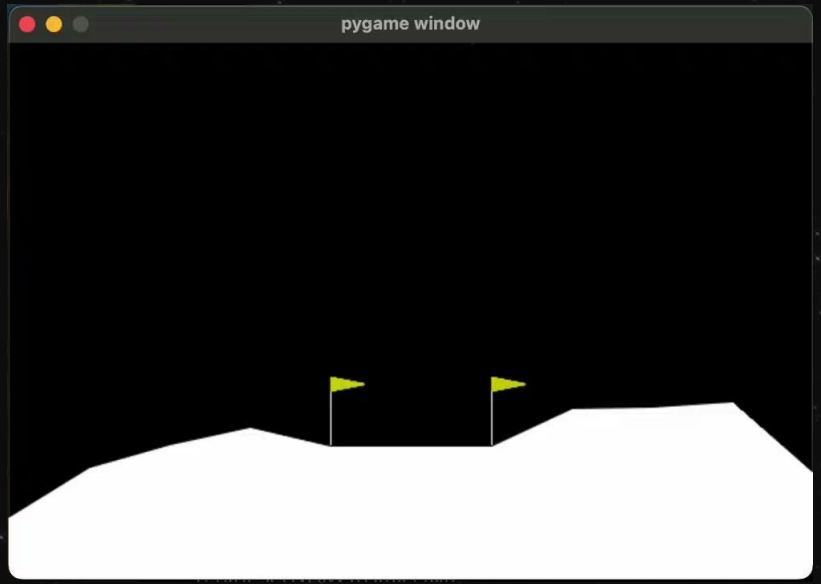
Neural Network Structure



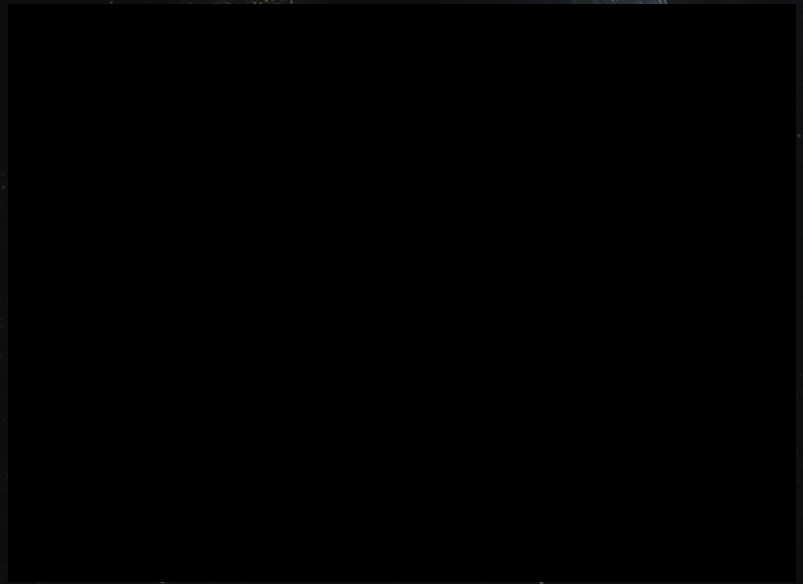
Rewards

- Measure the points per episode.
  - +100 to +300 for landing successfully between the flags (centered higher)
  - +10 for each leg making contact with the ground
  - - 0.3 points for each step taken
  - A crash is -100

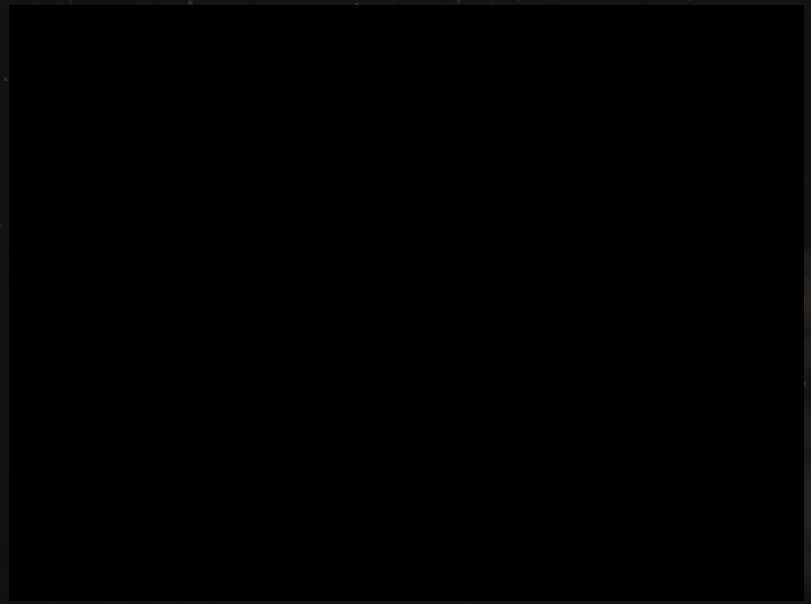
Starting completely Random



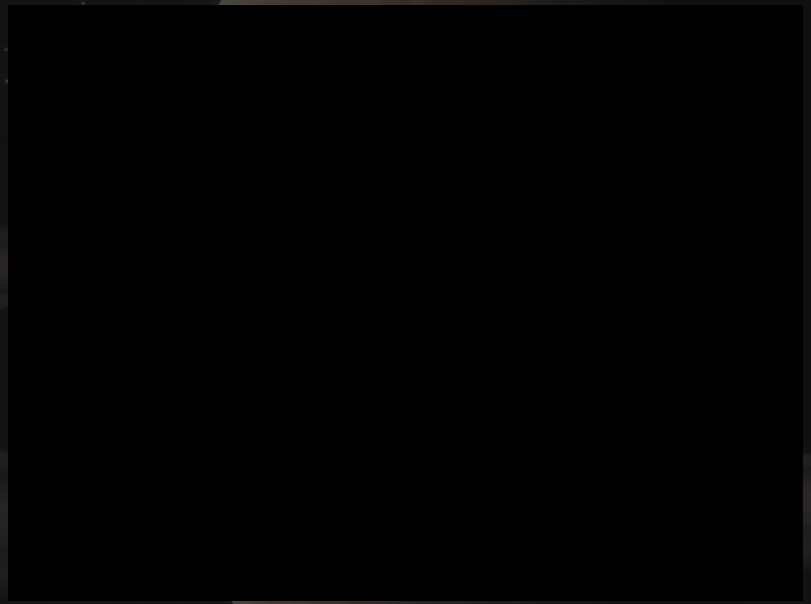
10 episodes



100 episodes

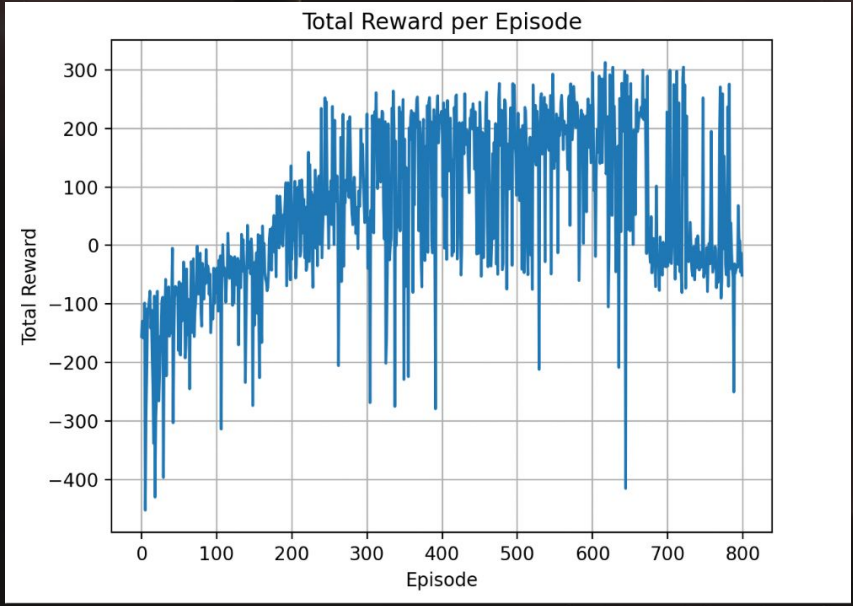


10 episodes

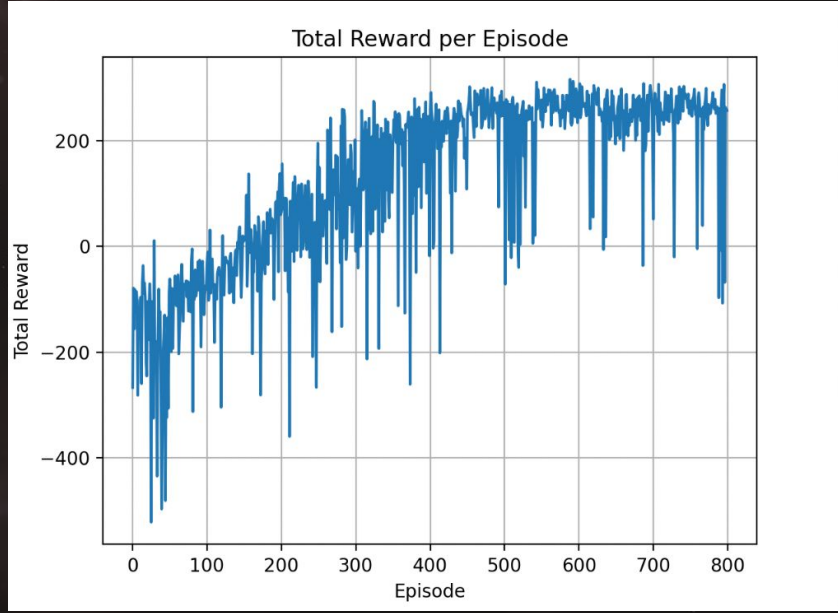


100 episodes

Issues We faced

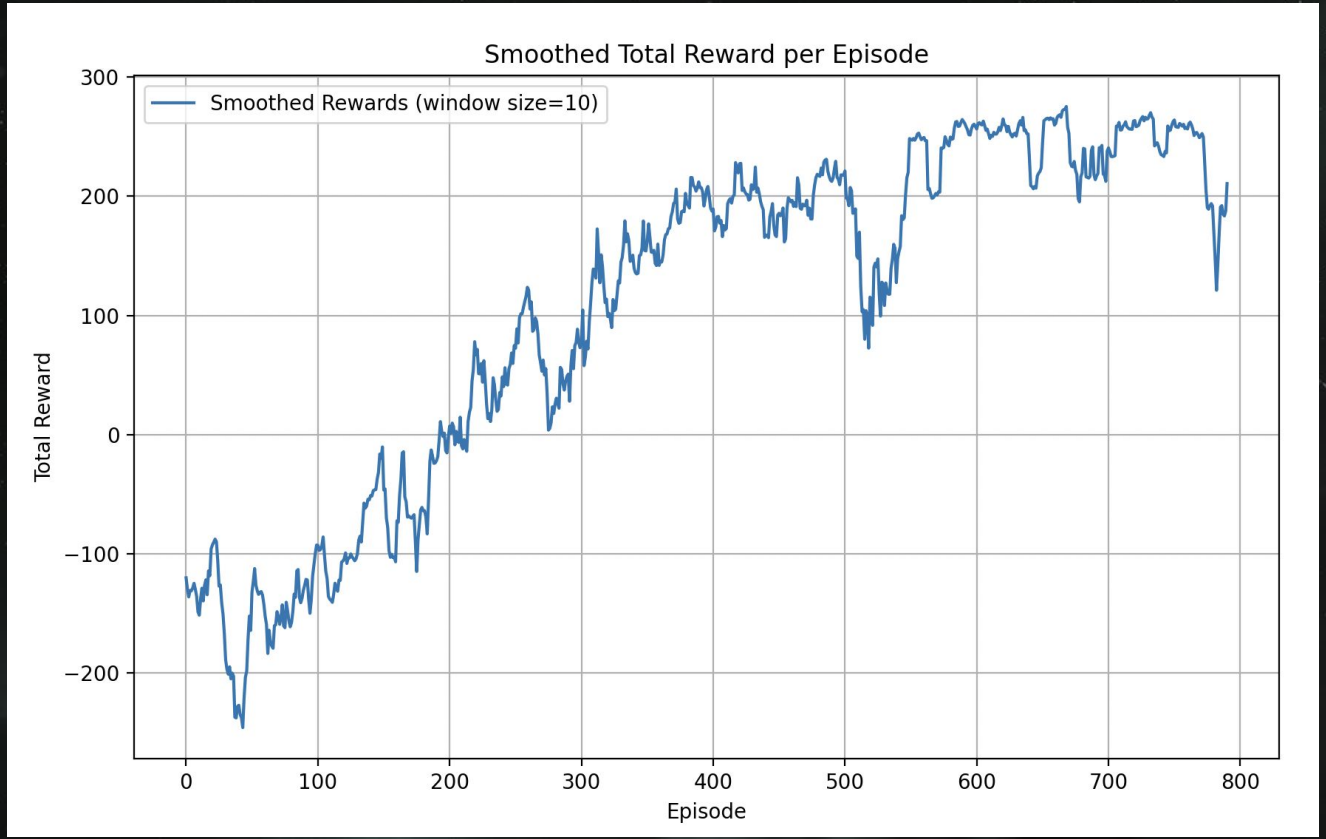


Update Frequency = 10



Update Frequency = 5

Results and analysis



DEMO

Why is this important?

Think about this application deployed on a real scale

Thanks!
Q&A