Double Deep Q-Learning

```
!apt-get update && apt-get install -y xvfb
```

```
Get:1 https://cloud.r-project.org/bin/linux/ubuntu jammy-cran40/ InRelease [3,626 B]
Get:2 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86_64  InRelease [1,581 B]
Get:3 http://security.ubuntu.com/ubuntu jammy-security InRelease [129 kB]
Hit:4 http://archive.ubuntu.com/ubuntu jammy InRelease
Get:5 https://r2u.stat.illinois.edu/ubuntu jammy InRelease [6,555 B]
Get:6 http://archive.ubuntu.com/ubuntu jammy-updates InRelease [128 kB]
Hit:7 https://ppa.launchpadcontent.net/deadsnakes/ppa/ubuntu jammy InRelease
Hit:8 https://ppa.launchpadcontent.net/graphics-drivers/ppa/ubuntu jammy InRelease
Get:9 http://archive.ubuntu.com/ubuntu jammy-backports InRelease [127 kB]
Hit:10 https://ppa.launchpadcontent.net/ubuntugis/ppa/ubuntu jammy InRelease
Get:11 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu2204/x86_64  Packages [1,113 kB]
Get:12 https://r2u.stat.illinois.edu/ubuntu jammy/main all Packages [8,482 kB]
Get:13 https://r2u.stat.illinois.edu/ubuntu jammy/main amd64 Packages [2,613 kB]
Get:14 http://security.ubuntu.com/ubuntu jammy-security/universe amd64 Packages [1,165 kB]
Get:15 http://security.ubuntu.com/ubuntu jammy-security/main amd64 Packages [2,425 kB]
Get:16 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 Packages [2,704 kB]
Get:17 http://archive.ubuntu.com/ubuntu jammy-updates/restricted amd64 Packages [3,358 kB]
Get:18 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 Packages [1,453 kB]
Fetched 23.7 MB in 8s (2,912 kB/s)
Reading package lists... Done
W: Skipping acquire of configured file 'main/source/Sources' as repository 'https://r2u.stat.illinois.edu/ubuntu jammy InRelease'
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
The following additional packages will be installed:
  libfontenc1 libxfont2 libxkbfile1 x11-xkb-utils xfonts-base xfonts-encodings xfonts-utils
  xserver-common
The following NEW packages will be installed:
  libfontenc1 libxfont2 libxkbfile1 x11-xkb-utils xfonts-base xfonts-encodings xfonts-utils
  xserver-common xvfb
0 upgraded, 9 newly installed, 0 to remove and 60 not upgraded.
Need to get 7,815 kB of archives.
After this operation, 11.9 MB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu jammy/main amd64 libfontenc1 amd64 1:1.1.4-1build3 [14.7 kB]
Get:2 http://archive.ubuntu.com/ubuntu jammy/main amd64 libxfont2 amd64 1:2.0.5-1build1 [94.5 kB]
Get:3 http://archive.ubuntu.com/ubuntu jammy/main amd64 libxkbfile1 amd64 1:1.1.0-1build3 [71.8 kB]
Get:4 http://archive.ubuntu.com/ubuntu jammy/main amd64 x11-xkb-utils amd64 7.7+5build4 [172 kB]
Get:5 http://archive.ubuntu.com/ubuntu jammy/main amd64 xfonts-encodings all 1:1.0.5-0ubuntu2 [578 kB]
Get:6 http://archive.ubuntu.com/ubuntu jammy/main amd64 xfonts-utils amd64 1:7.7+6build2 [94.6 kB]
Get:7 http://archive.ubuntu.com/ubuntu jammy/main amd64 xfonts-base all 1:1.0.5 [5,896 kB]
Get:8 http://archive.ubuntu.com/ubuntu jammy-updates/main amd64 xserver-common all 2:21.1.4-2ubuntu1.7~22.04.12 [28.7 kB]
Get:9 http://archive.ubuntu.com/ubuntu jammy-updates/universe amd64 xvfb amd64 2:21.1.4-2ubuntu1.7~22.04.12 [864 kB]
Fetched 7,815 kB in 1s (7,694 kB/s)
Selecting previously unselected package libfontenc1:amd64.
(Reading database ... 123629 files and directories currently installed.)
Preparing to unpack .../0-libfontenc1_1%3a1.1.4-1build3_amd64.deb ...
Unpacking libfontenc1:amd64 (1:1.1.4-1build3) ...
Selecting previously unselected package libxfont2:amd64.
Preparing to unpack .../1-libxfont2_1%3a2.0.5-1build1_amd64.deb ...
Unpacking libxfont2:amd64 (1:2.0.5-1build1) ...
Selecting previously unselected package libxkbfile1:amd64.
Preparing to unpack .../2-libxkbfile1_1%3a1.1.0-1build3_amd64.deb ...
Unpacking libxkbfile1:amd64 (1:1.1.0-1build3) ...
Selecting previously unselected package x11-xkb-utils.
Preparing to unpack .../3-x11-xkb-utils_7.7+5build4_amd64.deb ...
Unpacking x11-xkb-utils (7.7+5build4) ...
Selecting previously unselected package xfonts-encodings.
```

```
!pip install swig
```

```
Collecting swig
  Downloading swig-4.2.1.post0-py2.py3-none-manylinux_2_5_x86_64.manylinux1_x86_64.whl.metadata (3.5 kB)
Downloading swig-4.2.1.post0-py2.py3-none-manylinux_2_5_x86_64.manylinux1_x86_64.whl (1.8 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 1.8/1.8 MB 36.3 MB/s eta 0:00:00
Installing collected packages: swig
Successfully installed swig-4.2.1.post0
```

```
!pip install gym[box2d]==0.23.1
```

```
Collecting gym==0.23.1 (from gym[box2d]==0.23.1)
  Downloading gym-0.23.1.tar.gz (626 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 626.2/626.2 kB 25.6 MB/s eta 0:00:00
  Installing build dependencies ... done
  Getting requirements to build wheel ... done
  Preparing metadata (pyproject.toml) ... done
Requirement already satisfied: numpy>=1.18.0 in /usr/local/lib/python3.10/dist-packages (from gym==0.23.1->gym[box2d]==0.23.1) (1.26
Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gym==0.23.1->gym[box2d]==0.23.1)
Requirement already satisfied: gym-notices>=0.0.4 in /usr/local/lib/python3.10/dist-packages (from gym==0.23.1->gym[box2d]==0.23.1)
Collecting box2d-py==2.3.5 (from gym[box2d]==0.23.1)
```

```
  Downloading box2d-py-2.3.5.tar.gz (374 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 374.4/374.4 kB 30.0 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Collecting pygame==2.1.0 (from gym[box2d]==0.23.1)
  Downloading pygame-2.1.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (9.5 kB)
Downloading pygame-2.1.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (18.3 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 18.3/18.3 MB 82.0 MB/s eta 0:00:00
Building wheels for collected packages: gym, box2d-py
  Building wheel for gym (pyproject.toml) ... done
  Created wheel for gym: filename=gym-0.23.1-py3-none-any.whl size=701350 sha256=02604514dd47b376c25ccdf5ee3c03af7c96e89cfc3107bc2bc
  Stored in directory: /root/.cache/pip/wheels/1a/00/fb/fe5cf2860fb9b7bc860e28f00095a1f42c7b726dd6f42d1acc
  Building wheel for box2d-py (setup.py) ... done
  Created wheel for box2d-py: filename=box2d_py-2.3.5-cp310-cp310-linux_x86_64.whl size=2376098 sha256=5d6ce3050e47e86eb183e1bb0f831
  Stored in directory: /root/.cache/pip/wheels/db/8f/6a/eaaadf056fba10a98d986f6dce954e6201ba3126926fc5ad9e
Successfully built gym box2d-py
Installing collected packages: box2d-py, pygame, gym
  Attempting uninstall: pygame
    Found existing installation: pygame 2.6.1
    Uninstalling pygame-2.6.1:
      Successfully uninstalled pygame-2.6.1
  Attempting uninstall: gym
    Found existing installation: gym 0.25.2
    Uninstalling gym-0.25.2:
      Successfully uninstalled gym-0.25.2
Successfully installed box2d-py-2.3.5 gym-0.23.1 pygame-2.1.0
```

```
!pip install pytorch-lightning
```

```
Collecting pytorch-lightning
  Downloading pytorch_lightning-2.4.0-py3-none-any.whl.metadata (21 kB)
Requirement already satisfied: torch>=2.1.0 in /usr/local/lib/python3.10/dist-packages (from pytorch-lightning) (2.5.1+cu121)
Requirement already satisfied: tqdm>=4.57.0 in /usr/local/lib/python3.10/dist-packages (from pytorch-lightning) (4.66.6)
Requirement already satisfied: PyYAML>=5.4 in /usr/local/lib/python3.10/dist-packages (from pytorch-lightning) (6.0.2)
Requirement already satisfied: fsspec>=2022.5.0 in /usr/local/lib/python3.10/dist-packages (from fsspec[http]>=2022.5.0->pytorch-lig
Collecting torchmetrics>=0.7.0 (from pytorch-lightning)
  Downloading torchmetrics-1.6.0-py3-none-any.whl.metadata (20 kB)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from pytorch-lightning) (24.2)
Requirement already satisfied: typing-extensions>=4.4.0 in /usr/local/lib/python3.10/dist-packages (from pytorch-lightning) (4.12.2
Collecting lightning-utilities>=0.10.0 (from pytorch-lightning)
  Downloading lightning_utilities-0.11.8-py3-none-any.whl.metadata (5.2 kB)
Requirement already satisfied: aiohttp!=4.0.0a0,!=4.0.0a1 in /usr/local/lib/python3.10/dist-packages (from fsspec[http]>=2022.5.0->p
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from lightning-utilities>=0.10.0->pytorch-ligh
Requirement already satisfied: filelock in /usr/local/lib/python3.10/dist-packages (from torch>=2.1.0->pytorch-lightning) (3.16.1)
Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-packages (from torch>=2.1.0->pytorch-lightning) (3.4.2)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from torch>=2.1.0->pytorch-lightning) (3.1.4)
Requirement already satisfied: sympy==1.13.1 in /usr/local/lib/python3.10/dist-packages (from torch>=2.1.0->pytorch-lightning) (1.13
Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from sympy==1.13.1->torch>=2.1.0->pyto
Requirement already satisfied: numpy>1.20.0 in /usr/local/lib/python3.10/dist-packages (from torchmetrics>=0.7.0->pytorch-lightning
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1-
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[htt
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fssp
Requirement already satisfied: yarl<2.0,>=1.12.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec
Requirement already satisfied: async-timeout<5.0,>=4.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1-
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2->torch>=2.1.0->pytorch-lightn
Requirement already satisfied: idna>=2.0 in /usr/local/lib/python3.10/dist-packages (from yarl<2.0,>=1.12.0->aiohttp!=4.0.0a0,!=4.0
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.10/dist-packages (from yarl<2.0,>=1.12.0->aiohttp!=4.0.0a0
Downloading pytorch_lightning-2.4.0-py3-none-any.whl (815 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 815.2/815.2 kB 39.6 MB/s eta 0:00:00
Downloading lightning_utilities-0.11.8-py3-none-any.whl (26 kB)
Downloading torchmetrics-1.6.0-py3-none-any.whl (926 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 926.4/926.4 kB 44.6 MB/s eta 0:00:00
Installing collected packages: lightning-utilities, torchmetrics, pytorch-lightning
Successfully installed lightning-utilities-0.11.8 pytorch-lightning-2.4.0 torchmetrics-1.6.0
```

```
!pip install pyvirtualdisplay
```

```
Collecting pyvirtualdisplay
  Downloading PyVirtualDisplay-3.0-py3-none-any.whl.metadata (943 bytes)
Downloading PyVirtualDisplay-3.0-py3-none-any.whl (15 kB)
Installing collected packages: pyvirtualdisplay
Successfully installed pyvirtualdisplay-3.0
```

Setup virtual display

```
from pyvirtualdisplay import Display
Display(visible=False, size=(1400, 900)).start()
```

```
<pyvirtualdisplay.display.Display at 0x7e0da97134c0>
```

Import the necessary code libraries

```
import copy
import gym
import torch
import random

import numpy as np
import torch.nn.functional as F

from collections import deque, namedtuple
from IPython.display import HTML
from base64 import b64encode

from torch import Tensor, nn
from torch.utils.data import DataLoader
from torch.utils.data.dataset import IterableDataset
from torch.optim import AdamW

from pytorch_lightning import LightningModule, Trainer

from pytorch_lightning.callbacks import EarlyStopping

from gym.wrappers import RecordVideo, RecordEpisodeStatistics, TimeLimit


device = 'cuda:0' if torch.cuda.is_available() else 'cpu'
num_gpus = torch.cuda.device_count()
```

```
def display_video(episode=0):
  video_file = open(f'/content/videos/rl-video-episode-{episode}.mp4', "r+b").read()
  video_url = f"data:video/mp4;base64,{b64encode(video_file).decode()}"
  return HTML(f"<video width=600 controls><source src='{video_url}'></video>")
```

Create the Deep Q-Network

```
class DQN(nn.Module):

  def __init__(self, hidden_size, obs_size, n_actions):
    super().__init__()
    self.net = nn.Sequential(
        nn.Linear(obs_size, hidden_size),
        nn.ReLU(),
        nn.Linear(hidden_size, hidden_size),
        nn.ReLU(),
        nn.Linear(hidden_size, n_actions)
    )

  def forward(self, x):
    return self.net(x.float())
```

Create the policy

```
def epsilon_greedy(state, env, net, epsilon=0.0):
  if np.random.random() < epsilon:
    action = env.action_space.sample()
  else:
    state = torch.tensor([state]).to(device)
    q_values = net(state)
    _, action = torch.max(q_values, dim=1)
    action = int(action.item())
  return action
```

Create the replay buffer

```
class ReplayBuffer:

  def __init__(self, capacity):
    self.buffer = deque(maxlen=capacity)

  def __len__(self):
    return len(self.buffer)
```

```
  def append(self, experience):
    self.buffer.append(experience)

  def sample(self, batch_size):
    return random.sample(self.buffer, batch_size)
```

```
class RLDataset(IterableDataset):

  def __init__(self, buffer, sample_size=400):
    self.buffer = buffer
    self.sample_size = sample_size

  def __iter__(self):
    for experience in self.buffer.sample(self.sample_size):
      yield experience
```

Create the environment

```
def create_environment(name):
  env = gym.make(name)
  env = TimeLimit(env, max_episode_steps=400)
  env = RecordVideo(env, video_folder='./videos', episode_trigger=lambda x: x % 50 == 0)
  env = RecordEpisodeStatistics(env)
  return env
```

```
env = create_environment('LunarLander-v2')

for episode in range(10):
  done = False
  env.reset()
  while not done:
    action = env.action_space.sample()
    _, _, done, _ = env.step(action)
```

```
/usr/local/lib/python3.10/dist-packages/gym/wrappers/monitoring/video_recorder.py:115: DeprecationWarning: WARN: `env.metadata["vide
  logger.deprecation(
/usr/local/lib/python3.10/dist-packages/pygame/pkgdata.py:27: DeprecationWarning: pkg_resources is deprecated as an API. See https:/
  from pkg_resources import resource_stream, resource_exists
/usr/local/lib/python3.10/dist-packages/pkg_resources/__init__.py:3154: DeprecationWarning: Deprecated call to `pkg_resources.declar
Implementing implicit namespace packages (as specified in PEP 420) is preferred to `pkg_resources.declare_namespace`. See https://s
  declare_namespace(pkg)
/usr/local/lib/python3.10/dist-packages/pkg_resources/__init__.py:3154: DeprecationWarning: Deprecated call to `pkg_resources.declar
Implementing implicit namespace packages (as specified in PEP 420) is preferred to `pkg_resources.declare_namespace`. See https://s
  declare_namespace(pkg)
/usr/local/lib/python3.10/dist-packages/pkg_resources/__init__.py:3154: DeprecationWarning: Deprecated call to `pkg_resources.declar
Implementing implicit namespace packages (as specified in PEP 420) is preferred to `pkg_resources.declare_namespace`. See https://s
  declare_namespace(pkg)
/usr/local/lib/python3.10/dist-packages/gym/wrappers/monitoring/video_recorder.py:341: DeprecationWarning: Use shutil.which instead
  if distutils.spawn.find_executable("avconv") is not None:
/usr/local/lib/python3.10/dist-packages/gym/wrappers/monitoring/video_recorder.py:421: DeprecationWarning: distutils Version classes
  if distutils.version.LooseVersion(
```

```
display_video(episode=0)
```

0:03 / 0:03

Create the Deep Q-Learning algorithm

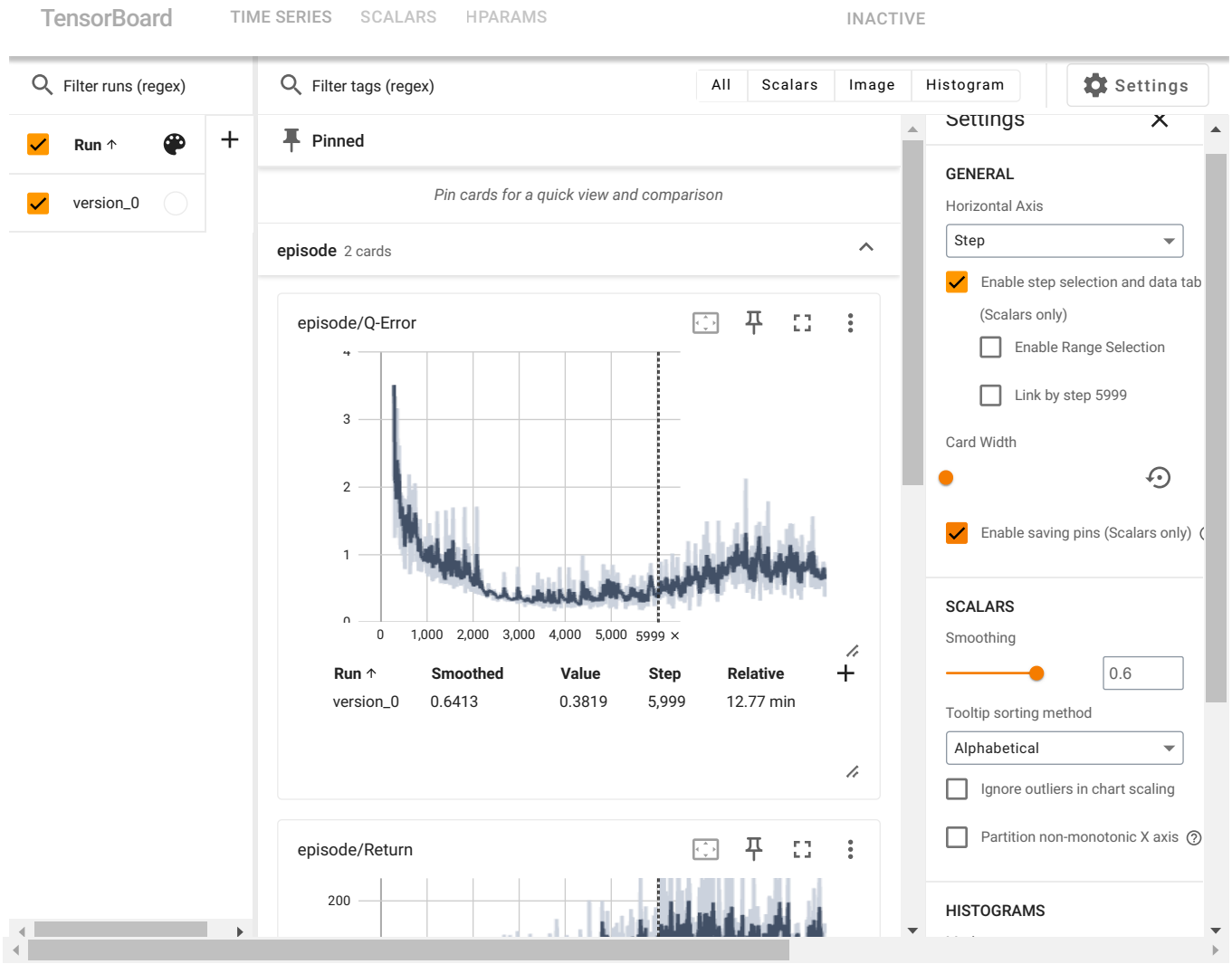
```
class DeepQLearning(LightningModule):

  # Initialize.
  def __init__(self, env_name, policy=epsilon_greedy, capacity=100_000,
               batch_size=256, lr=1e-3, hidden_size=128, gamma=0.99,
               loss_fn=F.smooth_l1_loss, optim=AdamW, eps_start=1.0, eps_end=0.15,
               eps_last_episode=100, samples_per_epoch=1_000, sync_rate=10):

    super().__init__()
    self.env = create_environment(env_name)

    obs_size = self.env.observation_space.shape[0]
    n_actions = self.env.action_space.n

    self.q_net = DQN(hidden_size, obs_size, n_actions)

    self.target_q_net = copy.deepcopy(self.q_net)

    self.policy = policy
    self.buffer = ReplayBuffer(capacity=capacity)

    self.save_hyperparameters()

    while len(self.buffer) < self.hparams.samples_per_epoch:
      print(f"{len(self.buffer)} samples in experience buffer. Filling...")
      self.play_episode(epsilon=self.hparams.eps_start)

  @torch.no_grad()
  def play_episode(self, policy=None, epsilon=0.):
    state = self.env.reset()
    done = False

    while not done:
      if policy:
        action = policy(state, self.env, self.q_net, epsilon=epsilon)
      else:
        action = self.env.action_space.sample()
      next_state, reward, done, info = self.env.step(action)
      exp = (state, action, reward, done, next_state)
      self.buffer.append(exp)
      state = next_state

  # Forward.
  def forward(self, x):
    return self.q_net(x)

  # Configure optimizers.
  def configure_optimizers(self):
    q_net_optimizer = self.hparams.optim(self.q_net.parameters(), lr=self.hparams.lr)
    return [q_net_optimizer]

  # Create dataloader.
  def train_dataloader(self):
    dataset = RLDataset(self.buffer, self.hparams.samples_per_epoch)
    dataloader = DataLoader(
        dataset=dataset,
        batch_size=self.hparams.batch_size
    )
    return dataloader

  # Training step.
  def training_step(self, batch, batch_idx):
    states, actions, rewards, dones, next_states = batch
    actions = actions.unsqueeze(1)
    rewards = rewards.unsqueeze(1)
    dones = dones.unsqueeze(1)

    state_action_values = self.q_net(states).gather(1, actions)

    with torch.no_grad():
      _, next_actions = self.q_net(next_states).max(dim=1, keepdim=True)
      next_action_values = self.target_q_net(next_states).gather(1, next_actions)
      next_action_values[dones] = 0.0

    expected_state_action_values = rewards + self.hparams.gamma * next_action_values

    loss = self.hparams.loss_fn(state_action_values, expected_state_action_values)
    self.log('episode/Q-Error', loss)
```

```
    return loss


  # Training epoch end.
  def on_train_epoch_end(self):

    epsilon = max(
        self.hparams.eps_end,
        self.hparams.eps_start - self.current_epoch / self.hparams.eps_last_episode
    )

    self.play_episode(policy=self.policy, epsilon=epsilon)
    self.log('episode/Return', self.env.return_queue[-1])

    if self.current_epoch % self.hparams.sync_rate == 0:
      self.target_q_net.load_state_dict(self.q_net.state_dict())
```

Purge logs and run the visualization tool (Tensorboard)

```
!rm -r /content/lightning_logs/
!rm -r /content/videos/
%load_ext tensorboard
%tensorboard --logdir /content/lightning_logs/
```

```
rm: cannot remove '/content/lightning_logs/': No such file or directory
```



Train the policy

```
import pytorch_lightning as pl
import warnings
warnings.filterwarnings('ignore')
```

```
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_c
  and should_run_async(code)
```

```
algo = DeepQLearning('LunarLander-v2')

trainer = pl.Trainer(
    accelerator="gpu" if num_gpus else "cpu", # Use 'gpu' if num_gpus is greater than 0, otherwise use 'cpu'
    devices=1, # Specify the number of GPUs or 'auto' for automatic detection
    max_epochs=1500,
    log_every_n_steps=10
)

trainer.fit(algo)
```

```
0 samples in experience buffer. Filling...
153 samples in experience buffer. Filling...
217 samples in experience buffer. Filling...
293 samples in experience buffer. Filling...
375 samples in experience buffer. Filling...
470 samples in experience buffer. Filling...
550 samples in experience buffer. Filling...
650 samples in experience buffer. Filling...
715 samples in experience buffer. Filling...
INFO:pytorch_lightning.utilities.rank_zero:GPU available: False, used: False
INFO:pytorch_lightning.utilities.rank_zero:TPU available: False, using: 0 TPU cores
INFO:pytorch_lightning.utilities.rank_zero:HPU available: False, using: 0 HPUs
854 samples in experience buffer. Filling...
952 samples in experience buffer. Filling...
INFO:pytorch_lightning.callbacks.model_summary:
  | Name         | Type | Params | Mode
-----------------------------------------------
0 | q_net        | DQN  | 18.2 K | train
1 | target_q_net | DQN  | 18.2 K | train
-----------------------------------------------
36.4 K    Trainable params
0         Non-trainable params
36.4 K    Total params
0.145     Total estimated model params size (MB)
14        Modules in train mode
0         Modules in eval mode
Epoch 1499:                                                                    4/? [00:00<00:00,  7.72it/s, v_num=0]
```

Check the resulting policy

```
display_video(episode=1450)
```


```
!zip -r /content/lightning_logs.zip /content/lightning_logs
```

```
  adding: content/lightning_logs/ (stored 0%)
  adding: content/lightning_logs/version_0/ (stored 0%)
  adding: content/lightning_logs/version_0/events.out.tfevents.1731951144.77a352ababc0.172.0 (deflated 68%)
  adding: content/lightning_logs/version_0/checkpoints/ (stored 0%)
  adding: content/lightning_logs/version_0/checkpoints/epoch=1499-step=6000.ckpt (deflated 13%)
  adding: content/lightning_logs/version_0/hparams.yaml (deflated 35%)
```

```
!zip -r /content/videos.zip /content/videos
```

```
  adding: content/videos/ (stored 0%)
  adding: content/videos/rl-video-episode-450.mp4 (deflated 12%)
  adding: content/videos/rl-video-episode-1400.meta.json (deflated 61%)
```

```
adding: content/videos/rl-video-episode-1200.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-600.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1000.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1500.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-550.mp4 (deflated 12%)
adding: content/videos/rl-video-episode-700.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-800.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-900.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-350.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1200.mp4 (deflated 15%)
adding: content/videos/rl-video-episode-300.mp4 (deflated 12%)
adding: content/videos/rl-video-episode-250.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-900.mp4 (deflated 13%)
adding: content/videos/rl-video-episode-500.mp4 (deflated 12%)
adding: content/videos/rl-video-episode-1250.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-550.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-700.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-1450.mp4 (deflated 13%)
adding: content/videos/rl-video-episode-1250.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1050.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-150.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1450.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-400.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-200.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1500.mp4 (deflated 15%)
adding: content/videos/rl-video-episode-850.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1050.mp4 (deflated 12%)
adding: content/videos/rl-video-episode-950.mp4 (deflated 13%)
adding: content/videos/rl-video-episode-600.mp4 (deflated 12%)
adding: content/videos/rl-video-episode-650.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-350.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-1350.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1150.mp4 (deflated 16%)
adding: content/videos/rl-video-episode-100.mp4 (deflated 13%)
adding: content/videos/rl-video-episode-1100.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1300.mp4 (deflated 15%)
adding: content/videos/rl-video-episode-450.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-50.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-200.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-1400.mp4 (deflated 15%)
adding: content/videos/rl-video-episode-1350.mp4 (deflated 15%)
adding: content/videos/rl-video-episode-1300.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-0.mp4 (deflated 14%)
adding: content/videos/rl-video-episode-250.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-500.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-750.mp4 (deflated 12%)
adding: content/videos/rl-video-episode-400.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-1000.mp4 (deflated 12%)
adding: content/videos/rl-video-episode-950.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1150.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-50.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-300.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-100.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-0.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-650.mp4 (deflated 12%)
```

Start coding or generate with AI.