

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

PyTorch Lightning

```
(Reading database \dots 123621 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.
     Preparing to unpack \dots/4-xfonts-encodings_1%3a1.0.5-0ubuntu2_all.deb \dots
     Unpacking xfonts-encodings (1:1.0.5-0ubuntu2) ..
     Selecting previously unselected package xfonts-utils.
     Preparing to unpack .../5-xfonts-utils_1%3a7.7+6build2_amd64.deb ...
     Unpacking xfonts-utils (1:7.7+6build2) ...
     Selecting previously unselected package xfonts-base.
     Preparing to unpack .../6-xfonts-base_1%3a1.0.5_all.deb ...
     Unpacking xfonts-base (1:1.0.5) ...
     Selecting previously unselected package xserver-common.
     Preparing to unpack .../7-xserver-common_2%3a21.1.4-2ubuntu1.7~22.04.11_all.deb ...
     Unpacking xserver-common (2:21.1.4-2ubuntu1.7~22.04.11) ...
     Selecting previously unselected package xvfb.
     Preparing to unpack .../8-xvfb_2%3a21.1.4-2ubuntu1.7~22.04.11_amd64.deb ...
     Unpacking xvfb (2:21.1.4-2ubuntu1.7~22.04.11) ...
     Setting up libfontenc1:amd64 (1:1.1.4-1build3) ...
     Setting up xfonts-encodings (1:1.0.5-0ubuntu2) ...
     Setting up libxkbfile1:amd64 (1:1.1.0-1build3) ...
     Setting up libxfont2:amd64 (1:2.0.5-1build1) ...
     Setting up x11-xkb-utils (7.7+5build4) ...
     Setting up xfonts-utils (1:7.7+6build2) ...
     Setting up xfonts-base (1:1.0.5) ...
     Setting up xserver-common (2:21.1.4-2ubuntu1.7~22.04.11) \dots
     Setting up xvfb (2:21.1.4-2ubuntu1.7~22.04.11) ...
     Processing triggers for man-db (2.10.2-1) ..
     Processing triggers for fontconfig (2.13.1-4.2ubuntu5) ...
     Processing triggers for libc-bin (2.35-0ubuntu3.4) ...
     /sbin/ldconfig.real: /usr/local/lib/libtbbmalloc_proxy.so.2 is not a symbolic link
     /sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_5.so.3 is not a symbolic link
     /sbin/ldconfig.real: /usr/local/lib/libur adapter opencl.so.0 is not a symbolic link
     /sbin/ldconfig.real: /usr/local/lib/libtbbmalloc.so.2 is not a symbolic link
     /sbin/ldconfig.real: /usr/local/lib/libtbbbind_2_0.so.3 is not a symbolic link
     /sbin/ldconfig.real: /usr/local/lib/libtbb.so.12 is not a symbolic link
     /sbin/ldconfig.real: /usr/local/lib/libur_adapter_level_zero.so.0 is not a symbolic link
     /sbin/ldconfig.real: /usr/local/lib/libur loader.so.0 is not a symbolic link
     /sbin/ldconfig.real: /usr/local/lib/libtbbbind.so.3 is not a symbolic link
!pip install swig

→ Collecting swig

       \label{lownloadingswig-4.2.1-py2.py3-none-manylinux2_5_x86_64.manylinux1_x86_64.whl.metadata~(3.6~kB)
     Downloading swig-4.2.1-py2.py3-none-manylinux_2_5_x86_64.manylinux1_x86_64.whl (1.9 MB)
                                                 - 1.9/1.9 MB 49.7 MB/s eta 0:00:00
     Installing collected packages: swig
     Successfully installed swig-4.2.1
!pip install pytorch_lightning
→ Collecting pytorch_lightning
```

Downloading pytorch_lightning-2.4.0-py3-none-any.whl.metadata (21 kB)

```
Pytorch Lightning Implementation DQN.ipynb - Colab
         Requirement already satisfied: torch>=2.1.0 in /usr/local/lib/python3.10/dist-packages (from pytorch_lightning) (2.4.1+cu121)
         Requirement already satisfied: tqdm>=4.57.0 in /usr/local/lib/python3.10/dist-packages (from pytorch_lightning) (4.66.5)
         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_liams from fsspec[http]>=2022.5.0->pytorch_liams fsspec[h
         Collecting torchmetrics>=0.7.0 (from pytorch_lightning)
             Downloading torchmetrics-1.4.3-py3-none-any.whl.metadata (19 kB)
         Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from pytorch lightning) (24.1)
         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-0.11.7-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->r
         Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities>=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-utilities=0.10.0->pytorch_lightning-u
         Requirement \ already \ satisfied: \ filelock \ in \ /usr/local/lib/python 3.10/dist-packages \ (from \ torch>=2.1.0->pytorch_lightning) \ (3.16.1)
         Requirement already satisfied: sympy in /usr/local/lib/python3.10/dist-packages (from torch>=2.1.0->pytorch_lightning) (1.13.3)
         Requirement already satisfied: networkx in /usr/local/lib/python3.10/dist-packages (from torch>=2.1.0->pytorch_lightning) (3.3)
         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: 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->fssr
         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_lightr
         Requirement already satisfied: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from sympy->torch>=2.1.0->pytorch_light
         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
         Downloading pytorch_lightning-2.4.0-py3-none-any.whl (815 kB)
                                                                                               815.2/815.2 kB 28.7 MB/s eta 0:00:00
         Downloading lightning utilities-0.11.7-py3-none-any.whl (26 kB)
         Downloading torchmetrics-1.4.3-py3-none-any.whl (869 kB)
                                                                                              - 869.5/869.5 kB <mark>39.7 MB/s</mark> eta 0:00:00
         Installing collected packages: lightning-utilities, torchmetrics, pytorch_lightning
         Successfully installed lightning-utilities-0.11.7 pytorch_lightning-2.4.0 torchmetrics-1.4.3
!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 23.3 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 26.8 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)
\label{lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_lower_low
                                                                                               18.3/18.3 MB 58.3 MB/s eta 0:00:00
Building wheels for collected packages: gym, box2d-py
    Building wheel for gym (pyproject.toml) \dots done
    Created wheel for gym: filename=gym-0.23.1-py3-none-any.whl size=701345 sha256=9370535fab7c9586bec41471aa26446a55995c62b6679408ec4
    Stored in directory: /root/.cache/pip/wheels/1a/00/fb/fe5cf2860fb9b7bc860e28f00095a1f42c7b726dd6f42d1acc
    Building wheel for box2d-py (setup.py) \dots done
    Created wheel for box2d-py: filename=box2d_py-2.3.5-cp310-cp310-linux_x86_64.whl size=2376102 sha256=fc631b42c113b4aea303db533f91c
    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 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 0x7a1f100899f0>

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()
#for saving videos in a video folder in mp4 file, which will read and write
#return the html too
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>")
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_
       and should_run_async(code)
```

Create the Deep Q-Network

```
#pytorch neural network,hidden_size-: no. of neurons
#obs_size-: vectors of each state
#n_actions-: the number of actions we can take
\hbox{\it\#prepare the resources from nn.module}
#will create list of layers
#apply the layers to the unput
#float tensors for giving to the neural 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

```
#policy-: state, env, neural network
#randomly selection for first condition else
```

```
# change state to tensor
#send this to device
#q_values from neural network
#select the higghest value with index as integer

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</pre>
```

Create the replay buffer

```
#will store observations and sample then with batch_size
#buffer will store the obsrvations of the states
#length of the self.buffer
#insert new info with append() method
#sample to take subset from 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)
#create RLDataset, so that pytorch can use
#sample and buffer
#going through one by one experience
#yield will experience, return the samples ony by one
class RLDataset(IterableDataset):
 def __init__(self, buffer, sample_size=400):
```

Create the environment

done = False
env.reset()
while not done:

action = env.action_space.sample()
_, _, done, _ = env.step(action)

vield experience

self.buffer = buffer

def __iter__(self):

self.sample_size = sample_size

for experience in self.buffer.sample(self.sample_size):

```
#name of env as parameter
# make env with gym.make()
#will recordvideo of the episodes and save as video for episode
#record episode stats from env

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

#sample running of env
env = create_environment('LunarLander-v2')

for episode in range(10):
```

4

```
/usr/local/lib/python3.10/dist-packages/gym/wrappers/monitoring/video_recorder.py:115: DeprecationWarning: WARN: `env.metadata["videogger.deprecation(
/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(
```

Create the Deep Q-Learning algorithm

```
#extend ligthning module class
#initialize, forward_pass, configure_optimizers, create data loader, training step,training_epoch_end
#env_name, policy-: epsilon-greedy, capacity-: for experience buffer, batch_size-: batch size, learning_rate
#hidden_size of neural network, discount factor, loss_function, optim,eps_start, eps_end, ep_last_episode,
#sample_per_epoch, sync_rate-:update target neural network
#call the init of super class
#create env
#create neural network
\#make target q net too with copy of nn
#policy
#replay buffer-:buffer
#save_hyperparameters-: keep in internal properties
#check len of buffer is lower than sample_per_epoch
#then we will fill the experience buffer
#play_episode method()-: action to be selected if policy is given
#exp will be appended in the buffer
#will update the state
#forward will compute the q_values
#will create a new optimizers by giving initial values
#train_dataloader-: will fetch samples from whole capacity and then batch_size will be selected and will be sent
#multiple batches will be loaded and trained
#traing_step: same as deep_q_learning
#training_epoch_end: take the epsilion value first
#cooling off value for 100 episodes
#then play the episode
#to make the target variable stable once every 10 epochs it will update.
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:</pre>
     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.
{\tt 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)
  next_action_values, _ = self.target_q_net(next_states).max(dim=1, keepdim=True)
 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):
  """Called at the end of the training epoch with the outputs of all training steps."""
 # The 'outputs' argument is now included
  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])
 print('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())
```

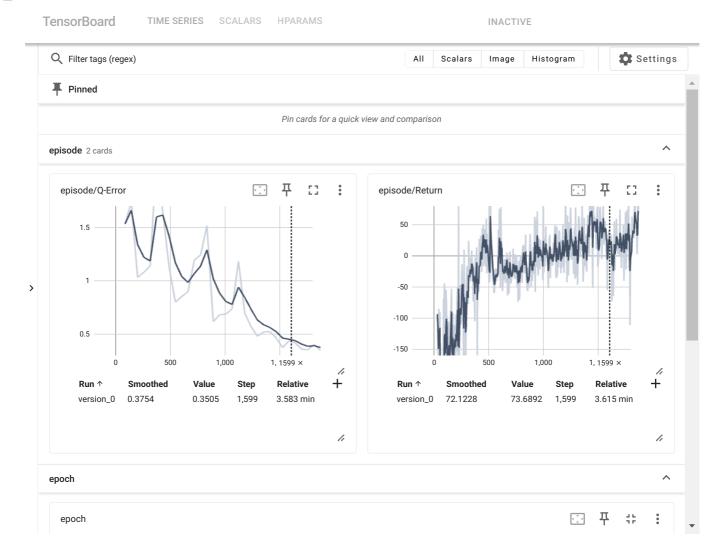
/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283: DeprecationWarning: `should_run_async` will not call `transform_c and should_run_async(code)

Purge logs and run the visualization tool (Tensorboard)

4

```
!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 p1
import warnings
warnings.filterwarnings('ignore')

algo = DeepQLearning('LunarLander-v2')

# Replace 'gpus=num_gpus' with 'accelerator="gpu"' or 'devices=num_gpus'

# If num_gpus is an integer representing the number of GPUs, use:
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=400,
    callbacks=[EarlyStopping(monitor='episode/Return', mode='max', patience=500)]

trainer.fit(algo)
```

Show hidden output

Check the resulting policy

display_video(episode=400)



0:13 / 0:13 !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.1728744153.a335900eee99.396.0 (deflated 68%) adding: content/lightning_logs/version_0/hparams.yaml (deflated 35%) adding: content/lightning_logs/version_0/checkpoints/ (stored 0%) adding: content/lightning_logs/version_0/checkpoints/epoch=399-step=1600.ckpt (deflated 13%) !zip -r /content/videos.zip /content/videos \rightarrow adding: content/videos/ (stored 0%) adding: content/videos/rl-video-episode-100.mp4 (deflated 10%) adding: content/videos/rl-video-episode-0.meta.json (deflated 61%) adding: content/videos/rl-video-episode-300.mp4 (deflated 11%) adding: content/videos/rl-video-episode-200.mp4 (deflated 11%) adding: content/videos/rl-video-episode-0.mp4 (deflated 13%) adding: content/videos/rl-video-episode-400.meta.json (deflated 61%) adding: content/videos/rl-video-episode-350.mp4 (deflated 12%) adding: content/videos/rl-video-episode-50.meta.json (deflated 61%) adding: content/videos/rl-video-episode-150.meta.json (deflated 61%) adding: content/videos/rl-video-episode-250.mp4 (deflated 12%) adding: content/videos/rl-video-episode-50.mp4 (deflated 12%) adding: content/videos/rl-video-episode-350.meta.json (deflated 61%) adding: content/videos/rl-video-episode-150.mp4 (deflated 11%)