∨ Hyperparameter tuning with Optuna


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

```
(Reading database ... 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 .../4-xfonts-encodings_1%3a1.0.5-0ubuntu2_all.deb ...
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) ...
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
  Downloading swig-4.2.1-py2.py3-none-manylinux_2_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 27.5 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)
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_lig
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->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: 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->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: mpmath<1.4,>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from sympy->torch>=2.1.0->pytorch_ligh
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 13.4 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 32.6 MB/s 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 12.5 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 20.2 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 34.1 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=701348 sha256=0e258ae0456452943d06f9bbbdc83c5358bb2539004c8d6b578
  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=2376100 sha256=b2635cf6bc06b0cbcf526258595ec
  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
```

```
!pip install optuna==2.7.0
```

```
Collecting optuna==2.7.0
  Downloading optuna-2.7.0-py3-none-any.whl.metadata (17 kB)
Collecting alembic (from optuna==2.7.0)
  Downloading alembic-1.13.3-py3-none-any.whl.metadata (7.4 kB)
Collecting cliff (from optuna==2.7.0)
  Downloading cliff-4.7.0-py3-none-any.whl.metadata (2.1 kB)
Collecting cmaes>=0.8.2 (from optuna==2.7.0)
  Downloading cmaes-0.11.1-py3-none-any.whl.metadata (18 kB)
Collecting colorlog (from optuna==2.7.0)
  Downloading colorlog-6.8.2-py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from optuna==2.7.0) (1.26.4)
Requirement already satisfied: packaging>=20.0 in /usr/local/lib/python3.10/dist-packages (from optuna==2.7.0) (24.1)
Requirement already satisfied: scipy!=1.4.0 in /usr/local/lib/python3.10/dist-packages (from optuna==2.7.0) (1.13.1)
Requirement already satisfied: sqlalchemy>=1.1.0 in /usr/local/lib/python3.10/dist-packages (from optuna==2.7.0) (2.0.35)
Requirement already satisfied: tqdm in /usr/local/lib/python3.10/dist-packages (from optuna==2.7.0) (4.66.5)
Requirement already satisfied: typing-extensions>=4.6.0 in /usr/local/lib/python3.10/dist-packages (from sqlalchemy>=1.1.0->optuna==
Requirement already satisfied: greenlet!=0.4.17 in /usr/local/lib/python3.10/dist-packages (from sqlalchemy>=1.1.0->optuna==2.7.0)
Collecting Mako (from alembic->optuna==2.7.0)
  Downloading Mako-1.3.5-py3-none-any.whl.metadata (2.9 kB)
Requirement already satisfied: PrettyTable>=0.7.2 in /usr/local/lib/python3.10/dist-packages (from cliff->optuna==2.7.0) (3.11.0)
Requirement already satisfied: PyYAML>=3.12 in /usr/local/lib/python3.10/dist-packages (from cliff->optuna==2.7.0) (6.0.2)
Collecting autopage>=0.4.0 (from cliff->optuna==2.7.0)
  Downloading autopage-0.5.2-py3-none-any.whl.metadata (7.9 kB)
Collecting cmd2>=1.0.0 (from cliff->optuna==2.7.0)
  Downloading cmd2-2.4.3-py3-none-any.whl.metadata (12 kB)
Collecting stevedore>=2.0.1 (from cliff->optuna==2.7.0)
  Downloading stevedore-5.3.0-py3-none-any.whl.metadata (2.3 kB)
Requirement already satisfied: attrs>=16.3.0 in /usr/local/lib/python3.10/dist-packages (from cmd2>=1.0.0->cliff->optuna==2.7.0) (24
Requirement already satisfied: pyperclip>=1.6 in /usr/local/lib/python3.10/dist-packages (from cmd2>=1.0.0->cliff->optuna==2.7.0) (1
Requirement already satisfied: wcwidth>=0.1.7 in /usr/local/lib/python3.10/dist-packages (from cmd2>=1.0.0->cliff->optuna==2.7.0) (0
Collecting pbr>=2.0.0 (from stevedore>=2.0.1->cliff->optuna==2.7.0)
  Downloading pbr-6.1.0-py2.py3-none-any.whl.metadata (3.4 kB)
Requirement already satisfied: MarkupSafe>=0.9.2 in /usr/local/lib/python3.10/dist-packages (from Mako->alembic->optuna==2.7.0) (2.1
Downloading optuna-2.7.0-py3-none-any.whl (293 kB)
                                  ━━━━━━━━━━━━━━━━━━━━━━━━━ 293.5/293.5 kB 11.6 MB/s eta 0:00:00
Downloading cmaes-0.11.1-py3-none-any.whl (35 kB)
Downloading alembic-1.13.3-py3-none-any.whl (233 kB)
                                  ━━━━━━━━━━━━━━━━━━━━━━━━━ 233.2/233.2 kB 18.9 MB/s eta 0:00:00
Downloading cliff-4.7.0-py3-none-any.whl (81 kB)
                                  ━━━━━━━━━━━━━━━━━━━━━━━━━ 81.2/81.2 kB 5.0 MB/s eta 0:00:00
Downloading colorlog-6.8.2-py3-none-any.whl (11 kB)
Downloading autopage-0.5.2-py3-none-any.whl (30 kB)
Downloading cmd2-2.4.3-py3-none-any.whl (147 kB)
                                  ━━━━━━━━━━━━━━━━━━━━━━━━━ 147.2/147.2 kB 11.2 MB/s eta 0:00:00
Downloading stevedore-5.3.0-py3-none-any.whl (49 kB)
                                  ━━━━━━━━━━━━━━━━━━━━━━━━━ 49.7/49.7 kB 4.1 MB/s eta 0:00:00
Downloading Mako-1.3.5-py3-none-any.whl (78 kB)
                                  ━━━━━━━━━━━━━━━━━━━━━━━━━ 78.6/78.6 kB 6.6 MB/s eta 0:00:00
Downloading pbr-6.1.0-py2.py3-none-any.whl (108 kB)
                                  ━━━━━━━━━━━━━━━━━━━━━━━━━ 108.5/108.5 kB 5.5 MB/s eta 0:00:00
Installing collected packages: pbr, Mako, colorlog, cmd2, cmaes, autopage, stevedore, alembic, cliff, optuna
Successfully installed Mako-1.3.5 alembic-1.13.3 autopage-0.5.2 cliff-4.7.0 cmaes-0.11.1 cmd2-2.4.3 colorlog-6.8.2 optuna-2.7.0 pbr
```

Setup virtual display

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

```
<pyvirtualdisplay.display.Display at 0x7c2c21b1dde0>
```

Import the necessary code libraries

```
import copy
import statistics
import gym
import torch
import optuna

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 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 gym.wrappers import RecordVideo, RecordEpisodeStatistics

from optuna.integration import PyTorchLightningPruningCallback

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

# Type redeclaration for backwards compatibility.
np.bool = np.bool_
```

⮂ /usr/local/lib/python3.10/dist-packages/optuna/storages/_rdb/models.py:35: MovedIn20Warning: The ``declarative_base()`` function is
    BaseModel: Any = declarative_base()
  /usr/local/lib/python3.10/dist-packages/optuna/visualization/matplotlib/_param_importances.py:33: MatplotlibDeprecationWarning: The
    cmap = cm.get_cmap("tab20c")

◀ ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬   ▶

```
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_c
    and should_run_async(code)

◀ ▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬▬   ▶

⌄ 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

```
Experience = namedtuple(
    "Experience",
    field_names=["state", "action", "reward", "done", "new_state"],
)
```

```
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):
```

```
        indices = np.random.choice(len(self.buffer), batch_size, replace=False)
        states, actions, rewards, dones, next_states = zip(*(self.buffer[idx] for idx in indices))

        return (
            np.array(states),
            np.array(actions),
            np.array(rewards, dtype=np.float32),
            np.array(dones, dtype=np.bool),
            np.array(next_states),
        )
```

```
class RLDataset(IterableDataset):

    def __init__(self, buffer, sample_size=200):
        self.buffer = buffer
        self.sample_size = sample_size

    def __iter__(self):
        states, actions, rewards, dones, new_states = self.buffer.sample(self.sample_size)
        for i in range(len(dones)):
            yield states[i], actions[i], rewards[i], dones[i], new_states[i]
```

⌄ Create the environment

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

⌄ Create the test/sampling function

```
@torch.no_grad()
def play_episode(env, q_net, buffer, policy=None, epsilon=0.):
    obs = env.reset()
    done = False

    while not done:
      if policy:
        action = policy(obs, env, q_net, epsilon=epsilon)
      else:
        action = env.action_space.sample()

      next_obs, reward, done, info = env.step(action)
      exp = Experience(obs, action, reward, done, next_obs)
      buffer.append(exp)
      obs = next_obs
```

⌄ Create the Deep Q-Learning algorithm

```
class DeepQLearning(LightningModule):

  def __init__(self, env_name, policy=epsilon_greedy, sample_fn=play_episode,
               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=10_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).to(device)
    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.hparams.sample_fn(self.env, self.q_net, self.buffer, epsilon=self.hparams.eps_start)
```

```python
    def forward(self, x):
      output = self.q_net(x)
      return output

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

    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

    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_state_values, _ = self.target_q_net(next_states).max(dim=1, keepdim=True)
        next_state_values[dones] = 0.0

      expected_state_action_values = rewards + self.hparams.gamma * next_state_values

      loss = self.hparams.loss_fn(state_action_values, expected_state_action_values)
      self.log('episode/MSE Loss', loss, on_step=False, on_epoch=True)
      return loss

  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.hparams.sample_fn(self.env, self.q_net, self.buffer, policy=self.policy, epsilon=epsilon)

      self.log("episode/Return", self.env.return_queue[-1])
      if self.current_epoch == 199:
        returns = list(self.env.return_queue)
        self.log("hp_metric", statistics.mean(returns))


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

> ⌄ Create the objective function

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

def objective(trial):
  lr = trial.suggest_float("lr", 1e-5, 1e-1, log=True)
  gamma = trial.suggest_float("gamma", 0.0, 1.0)
  hidden_size = trial.suggest_categorical("hidden_size", [32, 64, 128, 256])
  eps_end = trial.suggest_float("eps_end", 0.0, 0.3)

  algo = DeepQLearning(
      'LunarLander-v2',
      lr=lr,
      gamma=gamma,
      hidden_size=hidden_size,
      eps_end=eps_end
  )

  callback = PyTorchLightningPruningCallback(trial, monitor="episode/Return")

  trainer = pl.Trainer(
      accelerator="gpu" if num_gpus else "cpu", # Use 'gpu' if num_gpus is greater than 0, otherwise use 'cpu'
```

```
        devices=1,
        max_epochs=200,
        callbacks=[callback]
    )

    hyperparameters = dict(
        lr=lr,
        gamma=gamma,
        hidden_size=hidden_size,
        eps_end=eps_end
    )

    trainer.logger.log_hyperparams(hyperparameters)
    trainer.fit(algo)

    return trainer.callback_metrics["episode/Return"].item()
```

⌄ Create the optimization study

```
pruner = optuna.pruners.SuccessiveHalvingPruner()
study = optuna.create_study(direction="maximize", pruner=pruner)
```

```
[I 2024-10-12 16:11:22,010] A new study created in memory with name: no-name-83dc3347-6dfd-4a8c-8ac4-d6e5a7a80eea
```

⌄ Purge logs and run the visualization tool (Tensorboard)

```
# Start 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
rm: cannot remove '/content/videos/': No such file or directory
```

∨ Run the hyperparameter search

```
study.optimize(objective, n_trials=5)
```

⮒ **Show hidden output**

∨ Select and use the best hyperparameters

```
study.best_params
```

⮒ {'lr': 0.000862100613199551,
   'gamma': 0.6745711167247138,
   'hidden_size': 128,
   'eps_end': 0.23576196214250966}

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

⮒ The tensorboard extension is already loaded. To reload it, use:
    %reload_ext tensorboard
   Reusing TensorBoard on port 6006 (pid 2330), started 0:23:43 ago. (Use '!kill 2330' to kill it.)

TensorBoard   TIME SERIES   SCALARS   HPARAMS   INACTIVE

🔍 Filter   🔍 Filter tags (regex)   All   Scalars   Image   Histogram   ⚙ Settings

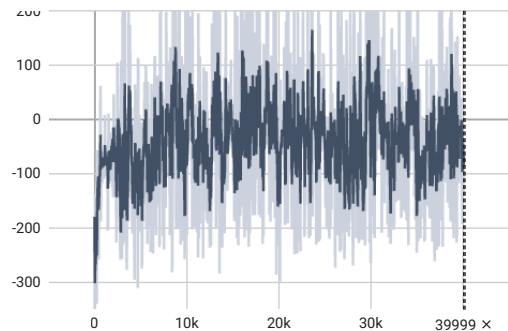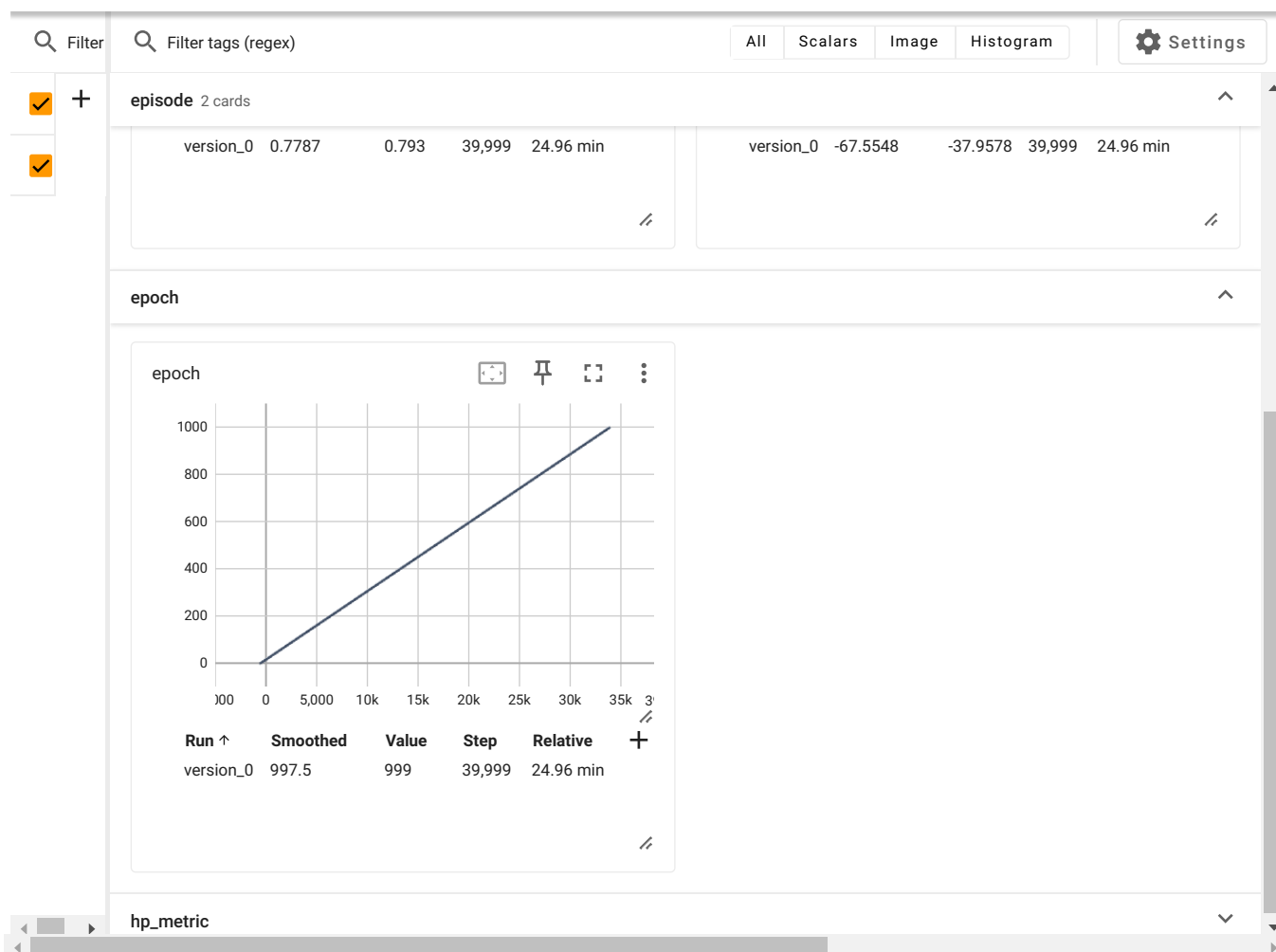☑ +   **episode** 2 cards   ∧

☑   version_0   0.7787   0.793   39,999   24.96 min   version_0   -67.5548   -37.9578 39,999 24.96 min

      epoch   ∧

      epoch   ⟦ ⟧ ☰ ⟦ ⟧ ⋮

      1000
       800
       600
       400
       200
         0
           00 0   5,000 10k 15k 20k 25k 30k 35k 3(
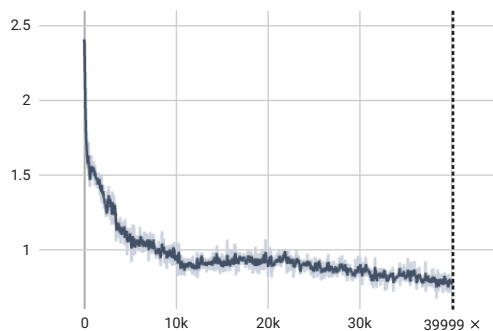
      Run ↑   **Smoothed**   **Value**   **Step**   **Relative**   +
      version_0   997.5   999   39,999   24.96 min

      hp_metric   ∨

◄ ▮▮ ►

```
algo = DeepQLearning('LunarLander-v2', **study.best_params)

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=1000
)

trainer.fit(algo)
```

```
0 samples in experience buffer. Filling...
136 samples in experience buffer. Filling...
210 samples in experience buffer. Filling...
311 samples in experience buffer. Filling...
384 samples in experience buffer. Filling...
452 samples in experience buffer. Filling...
545 samples in experience buffer. Filling...
622 samples in experience buffer. Filling...
756 samples in experience buffer. Filling...
824 samples in experience buffer. Filling...
904 samples in experience buffer. Filling...
986 samples in experience buffer. Filling...
1076 samples in experience buffer. Filling...
1179 samples in experience buffer. Filling...
1273 samples in experience buffer. Filling...
1375 samples in experience buffer. Filling...
1488 samples in experience buffer. Filling...
1549 samples in experience buffer. Filling...
1618 samples in experience buffer. Filling...
1717 samples in experience buffer. Filling...
1835 samples in experience buffer. Filling...
1892 samples in experience buffer. Filling...
1977 samples in experience buffer. Filling...
2073 samples in experience buffer. Filling...
2140 samples in experience buffer. Filling...
2226 samples in experience buffer. Filling...
2300 samples in experience buffer. Filling...
2397 samples in experience buffer. Filling...
2538 samples in experience buffer. Filling...
2618 samples in experience buffer. Filling...
2706 samples in experience buffer. Filling...
2802 samples in experience buffer. Filling...
2874 samples in experience buffer. Filling...
2984 samples in experience buffer. Filling...
3073 samples in experience buffer. Filling...
3177 samples in experience buffer. Filling...
3264 samples in experience buffer. Filling...
3333 samples in experience buffer. Filling...
3424 samples in experience buffer. Filling...
3520 samples in experience buffer. Filling...
3584 samples in experience buffer. Filling...
3659 samples in experience buffer. Filling...
3718 samples in experience buffer. Filling...
3814 samples in experience buffer. Filling...
3908 samples in experience buffer. Filling...
3970 samples in experience buffer. Filling...
4038 samples in experience buffer. Filling...
4160 samples in experience buffer. Filling...
4256 samples in experience buffer. Filling...
4319 samples in experience buffer. Filling...
4430 samples in experience buffer. Filling...
4524 samples in experience buffer. Filling...
4641 samples in experience buffer. Filling...
4754 samples in experience buffer. Filling...
4829 samples in experience buffer. Filling...
4899 samples in experience buffer. Filling...
4980 samples in experience buffer. Filling...
5117 samples in experience buffer. Filling...
5199 samples in experience buffer. Filling...
5277 samples in experience buffer. Filling...
5378 samples in experience buffer. Filling...
5445 samples in experience buffer. Filling...
5565 samples in experience buffer. Filling...
5633 samples in experience buffer. Filling...
5714 samples in experience buffer. Filling...
5843 samples in experience buffer. Filling...
5926 samples in experience buffer. Filling...
6030 samples in experience buffer. Filling...
6159 samples in experience buffer. Filling...
6271 samples in experience buffer. Filling...
6347 samples in experience buffer. Filling...
6455 samples in experience buffer. Filling...
6530 samples in experience buffer. Filling...
6629 samples in experience buffer. Filling...
6708 samples in experience buffer. Filling...
6791 samples in experience buffer. Filling...
6871 samples in experience buffer. Filling...
6943 samples in experience buffer. Filling...
7057 samples in experience buffer. Filling...
7202 samples in experience buffer. Filling...
7273 samples in experience buffer. Filling...
7384 samples in experience buffer. Filling...
7470 samples in experience buffer. Filling...
7581 samples in experience buffer. Filling...
7637 samples in experience buffer. Filling...
7707 samples in experience buffer. Filling...
7806 samples in experience buffer. Filling...
7867 samples in experience buffer. Filling...
7933 samples in experience buffer. Filling...
8018 samples in experience buffer. Filling...
```

```
8108 samples in experience buffer. Filling...
8176 samples in experience buffer. Filling...
8273 samples in experience buffer. Filling...
8418 samples in experience buffer. Filling...
8516 samples in experience buffer. Filling...
8636 samples in experience buffer. Filling...
8714 samples in experience buffer. Filling...
8822 samples in experience buffer. Filling...
8966 samples in experience buffer. Filling...
9079 samples in experience buffer. Filling...
9172 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
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
9312 samples in experience buffer. Filling...
9377 samples in experience buffer. Filling...
9450 samples in experience buffer. Filling...
9629 samples in experience buffer. Filling...
9716 samples in experience buffer. Filling...
9780 samples in experience buffer. Filling...
9845 samples in experience buffer. Filling...
9957 samples in experience buffer. Filling...

Epoch  999:                                                    40/? [00:00<00:00, 44.29it/s, v_num=0]
```

Check the resulting policy

```
display_video(episode=950)
```

0:08 / 0:08

```
!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/hparams.yaml (deflated 38%)
adding: content/lightning_logs/version_0/checkpoints/ (stored 0%)
adding: content/lightning_logs/version_0/checkpoints/epoch=999-step=40000.ckpt (deflated 10%)
adding: content/lightning_logs/version_0/events.out.tfevents.1728750919.54a00d33be76.470.5 (deflated 69%)
```

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

```
adding: content/videos/ (stored 0%)
adding: content/videos/rl-video-episode-100.mp4 (deflated 12%)
adding: content/videos/rl-video-episode-1000.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-900.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-0.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-850.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-550.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-700.mp4 (deflated 10%)
```

```
adding: content/videos/rl-video-episode-300.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-500.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-1100.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-200.mp4 (deflated 9%)
adding: content/videos/rl-video-episode-0.mp4 (deflated 14%)
adding: content/videos/rl-video-episode-550.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-400.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-350.mp4 (deflated 16%)
adding: content/videos/rl-video-episode-50.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-500.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-800.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-850.mp4 (deflated 21%)
adding: content/videos/rl-video-episode-950.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1100.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-700.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-600.mp4 (deflated 18%)
adding: content/videos/rl-video-episode-750.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-150.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1050.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-450.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-650.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-250.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-750.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-50.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-350.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-650.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-150.mp4 (deflated 11%)
adding: content/videos/rl-video-episode-100.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-800.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-400.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-1050.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-250.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-1000.mp4 (deflated 10%)
adding: content/videos/rl-video-episode-450.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-950.mp4 (deflated 9%)
adding: content/videos/rl-video-episode-300.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-900.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-600.meta.json (deflated 61%)
adding: content/videos/rl-video-episode-200.meta.json (deflated 61%)
```

Start coding or generate with AI.