# Generalized Advantage Estimation (GAE)

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

```
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%3a21.1.4-2ubuntu1.7~22.04.12_all.deb ...
Unpacking xserver-common (2:21.1.4-2ubuntu1.7~22.04.12) ...
Selecting previously unselected package xvfb.
Preparing to unpack .../8-xvfb_2%3a21.1.4-2ubuntu1.7~22.04.12_amd64.deb ...
Unpacking xvfb (2:21.1.4-2ubuntu1.7~22.04.12) ...
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 (2:21.1.4-2ubuntu1.7~22.04.12) ...
Setting up xvfb (2:21.1.4-2ubuntu1.7~22.04.12) ...
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/libtbb.so.12 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

/sbin/ldconfig.real: /usr/local/lib/libtcm.so.1 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libtbbmalloc_proxy.so.2 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/libtbbbind_2_0.so.3 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/libtcm_debug.so.1 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libhwloc.so.15 is not a symbolic link

/sbin/ldconfig.real: /usr/local/lib/libumf.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_5.so.3 is not a symbolic link
```

```
!pip install gym==0.23.1
```

```
Collecting gym==0.23.1
  Downloading gym-0.23.1.tar.gz (626 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 626.2/626.2 kB 12.7 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) (1.26.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gym==0.23.1) (3.1.0)
Requirement already satisfied: gym_notices>=0.0.4 in /usr/local/lib/python3.10/dist-packages (from gym==0.23.1) (0.0.8)
Building wheels for collected packages: gym
  Building wheel for gym (pyproject.toml) ... done
  Created wheel for gym: filename=gym-0.23.1-py3-none-any.whl size=701378 sha256=4ac33c89a79df0659674061aed22a4618dcb1ba3804e6dc757
  Stored in directory: /root/.cache/pip/wheels/1a/00/fb/fe5cf2860fb9b7bc860e28f00095a1f42c7b726dd6f42d1acc
Successfully built gym
Installing collected packages: gym
  Attempting uninstall: gym
    Found existing installation: gym 0.25.2
    Uninstalling gym-0.25.2:
      Successfully uninstalled gym-0.25.2
Successfully installed gym-0.23.1
```

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

```
Collecting pytorch-lightning
  Downloading pytorch_lightning-2.5.0-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) (4.67.1)
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-0.11.9-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: async-timeout<6.0,>=4.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1-
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: propcache>=0.2.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec
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,>=1.17.0->aiohttp!=4.0.0a0,!=4.0
Downloading pytorch_lightning-2.5.0-py3-none-any.whl (819 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 819.4/819.4 kB 18.4 MB/s eta 0:00:00
Downloading lightning_utilities-0.11.9-py3-none-any.whl (28 kB)
Downloading torchmetrics-1.6.0-py3-none-any.whl (926 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 926.4/926.4 kB 42.8 MB/s eta 0:00:00
Installing collected packages: lightning-utilities, torchmetrics, pytorch-lightning
Successfully installed lightning-utilities-0.11.9 pytorch-lightning-2.5.0 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
```

```
!pip install jax==0.3.14 jaxlib==0.3.14+cuda11.cudnn82 -f https://storage.googleapis.com/jax-releases/jax_cuda_releases.html
```

```
Looking in links: https://storage.googleapis.com/jax-releases/jax_cuda_releases.html
Collecting jax==0.3.14
  Downloading jax-0.3.14.tar.gz (990 kB)
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 990.1/990.1 kB 17.0 MB/s eta 0:00:00
  Preparing metadata (setup.py) ... done
Collecting jaxlib==0.3.14+cuda11.cudnn82
  Downloading https://storage.googleapis.com/jax-releases/cuda11/jaxlib-0.3.14%2Bcuda11.cudnn82-cp310-none-manylinux2014_x86_64.whl
     ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 161.9/161.9 MB 6.0 MB/s eta 0:00:00
Requirement already satisfied: absl-py in /usr/local/lib/python3.10/dist-packages (from jax==0.3.14) (1.4.0)
Requirement already satisfied: numpy>=1.19 in /usr/local/lib/python3.10/dist-packages (from jax==0.3.14) (1.26.4)
Requirement already satisfied: opt_einsum in /usr/local/lib/python3.10/dist-packages (from jax==0.3.14) (3.4.0)
Requirement already satisfied: scipy>=1.5 in /usr/local/lib/python3.10/dist-packages (from jax==0.3.14) (1.13.1)
Requirement already satisfied: typing_extensions in /usr/local/lib/python3.10/dist-packages (from jax==0.3.14) (4.12.2)
Requirement already satisfied: etils[epath] in /usr/local/lib/python3.10/dist-packages (from jax==0.3.14) (1.11.0)
Collecting flatbuffers<3.0,>=1.12 (from jaxlib==0.3.14+cuda11.cudnn82)
  Downloading flatbuffers-2.0.7-py2.py3-none-any.whl.metadata (872 bytes)
Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-packages (from etils[epath]->jax==0.3.14) (2024.10.0)
Requirement already satisfied: importlib_resources in /usr/local/lib/python3.10/dist-packages (from etils[epath]->jax==0.3.14) (6.4
Requirement already satisfied: zipp in /usr/local/lib/python3.10/dist-packages (from etils[epath]->jax==0.3.14) (3.21.0)
Downloading flatbuffers-2.0.7-py2.py3-none-any.whl (26 kB)
Building wheels for collected packages: jax
  Building wheel for jax (setup.py) ... done
  Created wheel for jax: filename=jax-0.3.14-py3-none-any.whl size=1147563 sha256=de04d7c7cff75a9dee470c34debcb09f5e7ad9d65a86ef981c
  Stored in directory: /root/.cache/pip/wheels/cc/b2/fd/3e8a8312ad916a677b3beb844bfba172999a08244fb1395149
Successfully built jax
Installing collected packages: flatbuffers, jaxlib, jax
  Attempting uninstall: flatbuffers
    Found existing installation: flatbuffers 24.3.25
    Uninstalling flatbuffers-24.3.25:
      Successfully uninstalled flatbuffers-24.3.25
  Attempting uninstall: jaxlib
    Found existing installation: jaxlib 0.4.33
    Uninstalling jaxlib-0.4.33:
      Successfully uninstalled jaxlib-0.4.33
```

```
    Attempting uninstall: jax
      Found existing installation: jax 0.4.33
      Uninstalling jax-0.4.33:
        Successfully uninstalled jax-0.4.33
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the sou
chex 0.1.88 requires jax>=0.4.27, but you have jax 0.3.14 which is incompatible.
chex 0.1.88 requires jaxlib>=0.4.27, but you have jaxlib 0.3.14+cuda11.cudnn82 which is incompatible.
flax 0.8.5 requires jax>=0.4.27, but you have jax 0.3.14 which is incompatible.
optax 0.2.4 requires jax>=0.4.27, but you have jax 0.3.14 which is incompatible.
optax 0.2.4 requires jaxlib>=0.4.27, but you have jaxlib 0.3.14+cuda11.cudnn82 which is incompatible.
orbax-checkpoint 0.6.4 requires jax>=0.4.26, but you have jax 0.3.14 which is incompatible.
tensorflow 2.17.1 requires flatbuffers>=24.3.25, but you have flatbuffers 2.0.7 which is incompatible.
Successfully installed flatbuffers-2.0.7 jax-0.3.14 jaxlib-0.3.14+cuda11.cudnn82
```

```
!pip install brax==0.1.1 #install compatible brax version

# Install other dependencies
!pip install gym==0.23.1
!pip install pytorch-lightning
!pip install pyvirtualdisplay
!pip install protobuf==3.20.3
```

```
Requirement already satisfied: brax==0.1.1 in /usr/local/lib/python3.10/dist-packages (0.1.1)
Requirement already satisfied: absl-py in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (1.4.0)
Requirement already satisfied: dataclasses in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (0.6)
Requirement already satisfied: dm-env in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (1.6)
Requirement already satisfied: flax in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (0.8.5)
Requirement already satisfied: grpcio in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (1.68.1)
Requirement already satisfied: gym in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (0.23.1)
Requirement already satisfied: jax in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (0.4.38)
Requirement already satisfied: jaxlib in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (0.4.38)
Requirement already satisfied: jaxopt in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (0.8.3)
Requirement already satisfied: jinja2 in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (3.1.4)
Requirement already satisfied: mujoco in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (3.2.6)
Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (1.26.4)
Requirement already satisfied: optax in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (0.2.4)
Requirement already satisfied: Pillow in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (11.0.0)
Requirement already satisfied: pytinyrenderer in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (0.0.14)
Requirement already satisfied: scipy in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (1.13.1)
Requirement already satisfied: tensorboardX in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (2.6.2.2)
Requirement already satisfied: trimesh==3.9.35 in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (3.9.35)
Requirement already satisfied: typing-extensions in /usr/local/lib/python3.10/dist-packages (from brax==0.1.1) (4.12.2)
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from trimesh==3.9.35->brax==0.1.1) (75.1.0)
Requirement already satisfied: dm-tree in /usr/local/lib/python3.10/dist-packages (from dm-env->brax==0.1.1) (0.1.8)
Requirement already satisfied: msgpack in /usr/local/lib/python3.10/dist-packages (from flax->brax==0.1.1) (1.1.0)
Requirement already satisfied: orbax-checkpoint in /usr/local/lib/python3.10/dist-packages (from flax->brax==0.1.1) (0.6.4)
Requirement already satisfied: tensorstore in /usr/local/lib/python3.10/dist-packages (from flax->brax==0.1.1) (0.1.71)
Requirement already satisfied: rich>=11.1 in /usr/local/lib/python3.10/dist-packages (from flax->brax==0.1.1) (13.9.4)
Requirement already satisfied: PyYAML>=5.4.1 in /usr/local/lib/python3.10/dist-packages (from flax->brax==0.1.1) (6.0.2)
Requirement already satisfied: ml_dtypes>=0.4.0 in /usr/local/lib/python3.10/dist-packages (from jax->brax==0.1.1) (0.4.1)
Requirement already satisfied: opt_einsum in /usr/local/lib/python3.10/dist-packages (from jax->brax==0.1.1) (3.4.0)
Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gym->brax==0.1.1) (3.1.0)
Requirement already satisfied: gym_notices>=0.0.4 in /usr/local/lib/python3.10/dist-packages (from gym->brax==0.1.1) (0.0.8)
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2->brax==0.1.1) (3.0.2)
Requirement already satisfied: etils[epath] in /usr/local/lib/python3.10/dist-packages (from mujoco->brax==0.1.1) (1.11.0)
Requirement already satisfied: glfw in /usr/local/lib/python3.10/dist-packages (from mujoco->brax==0.1.1) (2.8.0)
Requirement already satisfied: pyopengl in /usr/local/lib/python3.10/dist-packages (from mujoco->brax==0.1.1) (3.1.7)
Requirement already satisfied: chex>=0.1.87 in /usr/local/lib/python3.10/dist-packages (from optax->brax==0.1.1) (0.1.88)
Requirement already satisfied: packaging in /usr/local/lib/python3.10/dist-packages (from tensorboardX->brax==0.1.1) (24.2)
Requirement already satisfied: protobuf>=3.20 in /usr/local/lib/python3.10/dist-packages (from tensorboardX->brax==0.1.1) (4.25.5
Requirement already satisfied: toolz>=0.9.0 in /usr/local/lib/python3.10/dist-packages (from chex>=0.1.87->optax->brax==0.1.1) (0
Requirement already satisfied: markdown-it-py>=2.2.0 in /usr/local/lib/python3.10/dist-packages (from rich>=11.1->flax->brax==0.1
Requirement already satisfied: pygments<3.0.0,>=2.13.0 in /usr/local/lib/python3.10/dist-packages (from rich>=11.1->flax->brax==0
Requirement already satisfied: fsspec in /usr/local/lib/python3.10/dist-packages (from etils[epath]->mujoco->brax==0.1.1) (2024.1
Requirement already satisfied: importlib_resources in /usr/local/lib/python3.10/dist-packages (from etils[epath]->mujoco->brax==0
Requirement already satisfied: zipp in /usr/local/lib/python3.10/dist-packages (from etils[epath]->mujoco->brax==0.1.1) (3.21.0)
Requirement already satisfied: nest_asyncio in /usr/local/lib/python3.10/dist-packages (from orbax-checkpoint->flax->brax==0.1.1)
Requirement already satisfied: humanize in /usr/local/lib/python3.10/dist-packages (from orbax-checkpoint->flax->brax==0.1.1) (4.
Requirement already satisfied: mdurl~=0.1 in /usr/local/lib/python3.10/dist-packages (from markdown-it-py>=2.2.0->rich>=11.1->fla
Requirement already satisfied: gym==0.23.1 in /usr/local/lib/python3.10/dist-packages (0.23.1)
Requirement already satisfied: numpy>=1.18.0 in /usr/local/lib/python3.10/dist-packages (from gym==0.23.1) (1.26.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gym==0.23.1) (3.1.0)
Requirement already satisfied: gym_notices>=0.0.4 in /usr/local/lib/python3.10/dist-packages (from gym==0.23.1) (0.0.8)
Requirement already satisfied: pytorch-lightning in /usr/local/lib/python3.10/dist-packages (2.5.0)
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) (4.67.1)
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-
Requirement already satisfied: torchmetrics>=0.7.0 in /usr/local/lib/python3.10/dist-packages (from pytorch-lightning) (1.6.0)
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
Requirement already satisfied: lightning-utilities>=0.10.0 in /usr/local/lib/python3.10/dist-packages (from pytorch-lightning) (0
Requirement already satisfied: aiohttp!=4.0.0a0,!=4.0.0a1 in /usr/local/lib/python3.10/dist-packages (from fsspec[http]>=2022.5.0
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist-packages (from lightning-utilities>=0.10.0->pytorch-l
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
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->p
Requirement already satisfied: numpy>1.20.0 in /usr/local/lib/python3.10/dist-packages (from torchmetrics>=0.7.0->pytorch-lightni
Requirement already satisfied: aiohappyeyeballs>=2.3.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a
Requirement already satisfied: aiosignal>=1.1.2 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fssp
Requirement already satisfied: async-timeout<6.0,>=4.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a
Requirement already satisfied: attrs>=17.3.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fsspec[
Requirement already satisfied: frozenlist>=1.1.1 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fss
Requirement already satisfied: multidict<7.0,>=4.5 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->f
Requirement already satisfied: propcache>=0.2.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fssp
Requirement already satisfied: yarl<2.0,>=1.17.0 in /usr/local/lib/python3.10/dist-packages (from aiohttp!=4.0.0a0,!=4.0.0a1->fss
Requirement already satisfied: MarkupSafe>=2.0 in /usr/local/lib/python3.10/dist-packages (from jinja2->torch>=2.1.0->pytorch-lig
Requirement already satisfied: idna>=2.0 in /usr/local/lib/python3.10/dist-packages (from yarl<2.0,>=1.17.0->aiohttp!=4.0.0a0,!=4
Requirement already satisfied: pyvirtualdisplay in /usr/local/lib/python3.10/dist-packages (3.0)
Collecting protobuf==3.20.3
  Downloading protobuf-3.20.3-cp310-cp310-manylinux_2_12_x86_64.manylinux2010_x86_64.whl.metadata (679 bytes)
Downloading protobuf-3.20.3-cp310-cp310-manylinux_2_12_x86_64.manylinux2010_x86_64.whl (1.1 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 1.1/1.1 MB 18.6 MB/s eta 0:00:00
Installing collected packages: protobuf
  Attempting uninstall: protobuf
    Found existing installation: protobuf 4.25.5
    Uninstalling protobuf-4.25.5:
      Successfully uninstalled protobuf-4.25.5
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the
grpcio-status 1.62.3 requires protobuf>=4.21.6, but you have protobuf 3.20.3 which is incompatible.
```

```
tensorflow 2.17.1 requires flatbuffers>=24.3.25, but you have flatbuffers 2.0.7 which is incompatible.
Successfully installed protobuf-3.20.3
WARNING: The following packages were previously imported in this runtime:
  [google]
You must restart the runtime in order to use newly installed versions.
```

RESTART SESSION

```
!pip install jax==0.3.14 jaxlib==0.3.14+cuda11.cudnn82 -f https://storage.googleapis.com/jax-releases/jax_cuda_releases.html flax==0.4.2
```

```
  Downloading optax-0.2.2-py3-none-any.whl.metadata (8.1 kB)
INFO: pip is looking at multiple versions of chex to determine which version is compatible with other requirements. This could ta
Collecting chex>=0.1.86 (from optax->flax==0.4.2)
  Downloading chex-0.1.87-py3-none-any.whl.metadata (17 kB)
  Downloading chex-0.1.86-py3-none-any.whl.metadata (17 kB)
Collecting optax (from flax==0.4.2)
  Downloading optax-0.2.1-py3-none-any.whl.metadata (8.0 kB)
Collecting chex>=0.1.7 (from optax->flax==0.4.2)
  Downloading chex-0.1.85-py3-none-any.whl.metadata (17 kB)
  Downloading chex-0.1.84-py3-none-any.whl.metadata (17 kB)
INFO: pip is still looking at multiple versions of chex to determine which version is compatible with other requirements. This co
  Downloading chex-0.1.83-py3-none-any.whl.metadata (17 kB)
  Downloading chex-0.1.82-py3-none-any.whl.metadata (17 kB)
  Downloading chex-0.1.81-py3-none-any.whl.metadata (17 kB)
Requirement already satisfied: dm-tree>=0.1.5 in /usr/local/lib/python3.10/dist-packages (from chex>=0.1.7->optax->flax==0.4.2) (
  Downloading chex-0.1.7-py3-none-any.whl.metadata (17 kB)
Collecting optax (from flax==0.4.2)
  Downloading optax-0.2.0-py3-none-any.whl.metadata (8.0 kB)
INFO: This is taking longer than usual. You might need to provide the dependency resolver with stricter constraints to reduce run
INFO: pip is still looking at multiple versions of optax to determine which version is compatible with other requirements. This c
  Downloading optax-0.1.9-py3-none-any.whl.metadata (6.7 kB)
  Downloading optax-0.1.8-py3-none-any.whl.metadata (14 kB)
  Downloading optax-0.1.7-py3-none-any.whl.metadata (13 kB)
Collecting chex>=0.1.5 (from optax->flax==0.4.2)
  Downloading chex-0.1.6-py3-none-any.whl.metadata (17 kB)
Requirement already satisfied: toolz>=0.9.0 in /usr/local/lib/python3.10/dist-packages (from chex>=0.1.5->optax->flax==0.4.2) (0.
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.10/dist-packages (from python-dateutil>=2.7->matplotlib->flax==
Downloading flax-0.4.2-py3-none-any.whl (186 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 186.4/186.4 kB 7.3 MB/s eta 0:00:00
Downloading optax-0.1.7-py3-none-any.whl (154 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 154.1/154.1 kB 6.1 MB/s eta 0:00:00
Downloading chex-0.1.6-py3-none-any.whl (87 kB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 87.9/87.9 kB 3.9 MB/s eta 0:00:00
Installing collected packages: jaxlib, jax, chex, optax, flax
  Attempting uninstall: jaxlib
    Found existing installation: jaxlib 0.4.38
    Uninstalling jaxlib-0.4.38:
      Successfully uninstalled jaxlib-0.4.38
  Attempting uninstall: jax
    Found existing installation: jax 0.4.38
    Uninstalling jax-0.4.38:
      Successfully uninstalled jax-0.4.38
  Attempting uninstall: chex
    Found existing installation: chex 0.1.88
    Uninstalling chex-0.1.88:
      Successfully uninstalled chex-0.1.88
  Attempting uninstall: optax
    Found existing installation: optax 0.2.4
    Uninstalling optax-0.2.4:
      Successfully uninstalled optax-0.2.4
  Attempting uninstall: flax
    Found existing installation: flax 0.8.5
    Uninstalling flax-0.8.5:
      Successfully uninstalled flax-0.8.5
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the
orbax-checkpoint 0.6.4 requires jax>=0.4.26, but you have jax 0.3.14 which is incompatible.
Successfully installed chex-0.1.6 flax-0.4.2 jax-0.3.14 jaxlib-0.3.14+cuda11.cudnn82 optax-0.1.7
```

```
!pip install scipy==1.9.0
```

```
Collecting scipy==1.9.0
  Downloading scipy-1.9.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (2.2 kB)
Collecting numpy<1.25.0,>=1.18.5 (from scipy==1.9.0)
  Downloading numpy-1.24.4-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (5.6 kB)
Downloading scipy-1.9.0-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (43.9 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 43.9/43.9 MB 8.8 MB/s eta 0:00:00
Downloading numpy-1.24.4-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl (17.3 MB)
   ━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━━ 17.3/17.3 MB 35.9 MB/s eta 0:00:00
Installing collected packages: numpy, scipy
  Attempting uninstall: numpy
    Found existing installation: numpy 1.26.4
    Uninstalling numpy-1.26.4:
      Successfully uninstalled numpy-1.26.4
  Attempting uninstall: scipy
    Found existing installation: scipy 1.13.1
    Uninstalling scipy-1.13.1:
      Successfully uninstalled scipy-1.13.1
ERROR: pip's dependency resolver does not currently take into account all the packages that are installed. This behaviour is the sou
albumentations 1.4.20 requires scipy>=1.10.0, but you have scipy 1.9.0 which is incompatible.
orbax-checkpoint 0.6.4 requires jax>=0.4.26, but you have jax 0.3.14 which is incompatible.
pymc 5.19.1 requires numpy>=1.25.0, but you have numpy 1.24.4 which is incompatible.
scikit-image 0.25.0 requires scipy>=1.11.2, but you have scipy 1.9.0 which is incompatible.
tensorflow 2.17.1 requires flatbuffers>=24.3.25, but you have flatbuffers 2.0.7 which is incompatible.
Successfully installed numpy-1.24.4 scipy-1.9.0
WARNING: The following packages were previously imported in this runtime:
  [numpy]
You must restart the runtime in order to use newly installed versions.
```

RESTART SESSION

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

## Setup virtual display

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

```
<pyvirtualdisplay.display.Display at 0x7cfadce5be50>
```

## Import the necessary code libraries

```
import copy
import torch
import random
import gym
import matplotlib
import functools
import itertools
import math

import numpy as np
import matplotlib.pyplot as plt

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 torch.distributions import Normal

from pytorch_lightning import LightningModule, Trainer

import brax
from brax import envs
from brax.envs import to_torch
from brax.io import html

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

v = torch.ones(1, device='cpu')
```

```
/usr/local/lib/python3.10/dist-packages/trimesh/interfaces/scad.py:28: DeprecationWarning: Use shutil.which instead of find_executab
  _scad_executable = find_executable('openscad', path=_search_path)
/usr/local/lib/python3.10/dist-packages/trimesh/interfaces/blender.py:36: DeprecationWarning: Use shutil.which instead of find_execu
  _blender_executable = find_executable('blender', path=_search_path)
/usr/local/lib/python3.10/dist-packages/trimesh/interfaces/vhacd.py:21: DeprecationWarning: Use shutil.which instead of find_executa
  _vhacd_executable = find_executable(_name, path=_search_path)
/usr/local/lib/python3.10/dist-packages/trimesh/curvature.py:12: DeprecationWarning: Please use `coo_matrix` from the `scipy.sparse`
  from scipy.sparse.coo import coo_matrix
/usr/local/lib/python3.10/dist-packages/trimesh/exchange/ply.py:953: DeprecationWarning: Use shutil.which instead of find_executable
  draco_encoder = find_executable('draco_encoder')
/usr/local/lib/python3.10/dist-packages/trimesh/exchange/binvox.py:18: DeprecationWarning: Use shutil.which instead of find_executab
  binvox_encoder = find_executable('binvox')
/usr/local/lib/python3.10/dist-packages/trimesh/path/exchange/dxf.py:1101: DeprecationWarning: Use shutil.which instead of find_exec
  _teigha = find_executable(_name)
```

```
@torch.no_grad()
def create_video(env, episode_length, policy=None):
  qp_array = []
  state = env.reset()
  for i in range(episode_length):
    if policy:
      loc, scale = policy(state)
      sample = torch.normal(loc, scale)
      action = torch.tanh(sample)
    else:
      action = env.action_space.sample()
    state, _, _, _ = env.step(action)
    qp_array.append(env.unwrapped._state.qp)
  return HTML(html.render(env.unwrapped._env.sys, qp_array))


@torch.no_grad()
def test_agent(env, episode_length, policy, episodes=10):

  ep_returns = []
  for ep in range(episodes):
    state = env.reset()
    done = False
    ep_ret = 0.0

    while not done:
      loc, scale = policy(state)
      sample = torch.normal(loc, scale)
      action = torch.tanh(sample)
      state, reward, done, info = env.step(action)
      ep_ret += reward.item()

    ep_returns.append(ep_ret)

  return sum(ep_returns) / episodes
```

## Create the policy

```
class GradientPolicy(nn.Module):

  def __init__(self, in_features, out_dims, hidden_size=128):
    super().__init__()
    self.fc1 = nn.Linear(in_features, hidden_size)
    self.fc2 = nn.Linear(hidden_size, hidden_size)
    self.fc_mu = nn.Linear(hidden_size, out_dims)
    self.fc_std = nn.Linear(hidden_size, out_dims)

  def forward(self, x):
    x = F.relu(self.fc1(x))
    x = F.relu(self.fc2(x))
    loc = self.fc_mu(x)
    loc = torch.tanh(loc)
    scale = self.fc_std(x)
    scale = F.softplus(scale) + 0.001
    return loc, scale
```

## Create the value network

```
class ValueNet(nn.Module):

  def __init__(self, in_features, hidden_size=128):
    super().__init__()
    self.fc1 = nn.Linear(in_features, hidden_size)
```

```
    self.fc2 = nn.Linear(hidden_size, hidden_size)
    self.fc3 = nn.Linear(hidden_size, 1)

  def forward(self, x):
    x = F.relu(self.fc1(x))
    x = F.relu(self.fc2(x))
    x = self.fc3(x)
    return x
```

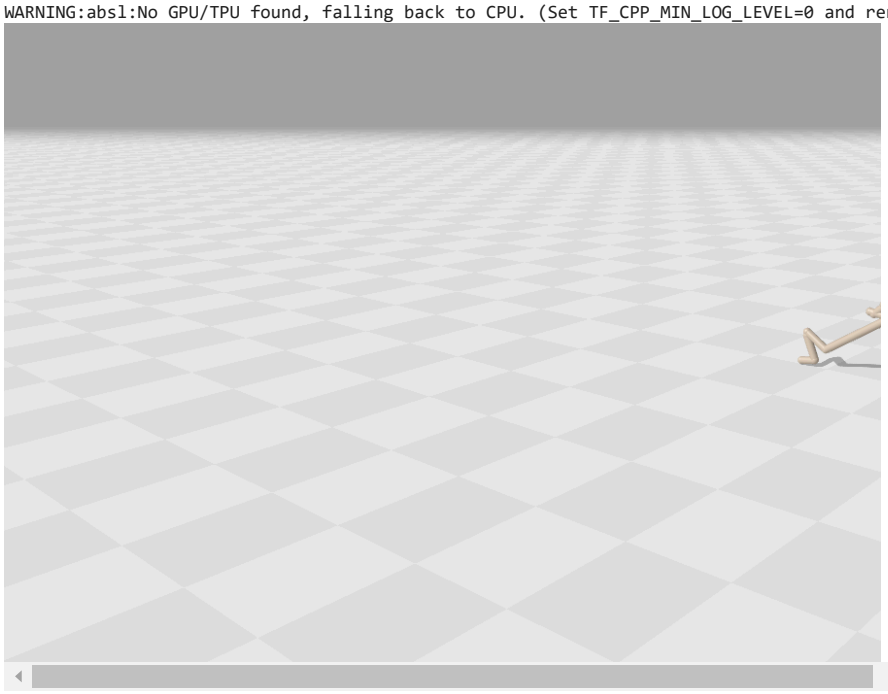
## Create the environment

```
class RunningMeanStd:
    # https://en.wikipedia.org/wiki/Algorithms_for_calculating_variance#Parallel_algorithm
    def __init__(self, epsilon=1e-4, shape=()):
        self.mean = torch.zeros(shape, dtype=torch.float32).to(device)
        self.var = torch.ones(shape, dtype=torch.float32).to(device)
        self.count = epsilon

    def update(self, x):
        batch_mean = torch.mean(x, dim=0)
        batch_var = torch.var(x, dim=0)
        batch_count = x.shape[0]
        self.update_from_moments(batch_mean, batch_var, batch_count)

    def update_from_moments(self, batch_mean, batch_var, batch_count):
        self.mean, self.var, self.count = update_mean_var_count_from_moments(
            self.mean, self.var, self.count, batch_mean, batch_var, batch_count
        )


def update_mean_var_count_from_moments(
    mean, var, count, batch_mean, batch_var, batch_count
):
    delta = batch_mean - mean
    tot_count = count + batch_count

    new_mean = mean + delta * batch_count / tot_count
    m_a = var * count
    m_b = batch_var * batch_count
    M2 = m_a + m_b + torch.square(delta) * count * batch_count / tot_count
    new_var = M2 / tot_count
    new_count = tot_count

    return new_mean, new_var, new_count


class NormalizeObservation(gym.core.Wrapper):

    def __init__(self, env, epsilon=1e-8):
        super().__init__(env)
        self.num_envs = getattr(env, "num_envs", 1)
        self.obs_rms = RunningMeanStd(shape=self.observation_space.shape[-1])
        self.epsilon = epsilon

    def step(self, action):
        obs, rews, dones, infos = self.env.step(action)
        obs = self.normalize(obs)
        return obs, rews, dones, infos

    def reset(self, **kwargs):
        return_info = kwargs.get("return_info", False)
        if return_info:
            obs, info = self.env.reset(**kwargs)
        else:
            obs = self.env.reset(**kwargs)
        obs = self.normalize(obs)
        if not return_info:
            return obs
        else:
            return obs, info

    def normalize(self, obs):
        self.obs_rms.update(obs)
        return (obs - self.obs_rms.mean) / torch.sqrt(self.obs_rms.var + self.epsilon)
```

```
entry_point = functools.partial(envs.create_gym_env, env_name='halfcheetah')
gym.register('brax-halfcheetah-v0', entry_point=entry_point)
```

```
env = gym.make("brax-halfcheetah-v0", episode_length=1000)
env = to_torch.JaxToTorchWrapper(env, device=device)
create_video(env, 1000)
```

```
WARNING:absl:No GPU/TPU found, falling back to CPU. (Set TF_CPP_MIN_LOG_LEVEL=0 and rerun for more i
```




```
def create_env(env_name, num_envs=256, episode_length=1000):
  env = gym.make(env_name, batch_size=num_envs, episode_length=episode_length)
  env = to_torch.JaxToTorchWrapper(env, device=device)
  env = NormalizeObservation(env)
  return env
```

```
env = create_env('brax-halfcheetah-v0', num_envs=10)
obs = env.reset()
print("Num envs: ", obs.shape[0], "Obs dimentions: ", obs.shape[1])
```

```
Num envs:  10 Obs dimentions:  18
```

```
env.action_space
```

```
Box(-1.0, 1.0, (10, 6), float32)
```

```
obs, reward, done, info = env.step(env.action_space.sample())
```

```
info.keys()
```

```
/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)
dict_keys(['first_obs', 'first_qp', 'reward_ctrl', 'reward_run', 'steps', 'truncation', 'x_position', 'x_velocity'])
```

## Create the dataset

```
class RLDataset(IterableDataset):

  def __init__(self, env, policy, value_net,
               samples_per_epoch, gamma, lamb, repeats):

    self.samples_per_epoch = samples_per_epoch
    self.gamma = gamma
    self.lamb = lamb
    self.repeats = repeats
    self.env = env
    self.policy = policy
    self.value_net = value_net
    self.obs = self.env.reset()

  @torch.no_grad()
  def __iter__(self):
    transitions = []
    for step in range(self.samples_per_epoch):
      loc, scale = self.policy(self.obs)
```

```
      action = torch.normal(loc, scale)
      next_obs, reward, done, info = self.env.step(action)
      transitions.append((self.obs, loc, scale, action, reward, done, next_obs))
      self.obs = next_obs

    transitions  = map(torch.stack, zip(*transitions))
    obs_b, loc_b, scale_b, action_b, reward_b, done_b, next_obs_b = transitions
    reward_b = reward_b.unsqueeze(dim=-1)
    done_b = done_b.unsqueeze(dim=-1)

    values_b = self.value_net(obs_b)
    next_values_b = self.value_net(next_obs_b)

    td_error_b = reward_b + (1 - done_b) * self.gamma * next_values_b - values_b

    running_gae = torch.zeros((self.env.num_envs, 1), dtype=torch.float32, device=device)
    gae_b = torch.zeros_like(td_error_b)

    for row in range(self.samples_per_epoch - 1, -1, -1):
      running_gae = td_error_b[row] + (1 - done_b[row]) * self.gamma * self.lamb * running_gae
      gae_b[row] = running_gae

    target_b = gae_b + values_b

    num_samples = self.samples_per_epoch * self.env.num_envs
    reshape_fn = lambda x: x.view(num_samples, -1)
    batch = [obs_b, loc_b, scale_b, action_b, reward_b, gae_b, target_b]

    obs_b, loc_b, scale_b, action_b, reward_b, gae_b, target_b = map(reshape_fn, batch)

    for repeat in range(self.repeats):
      idx = list(range(num_samples))
      random.shuffle(idx)

      for i in idx:
        yield obs_b[i], loc_b[i], scale_b[i], action_b[i], reward_b[i], gae_b[i], target_b[i]
```

## Create PPO with generalized advantage estimation (GAE)

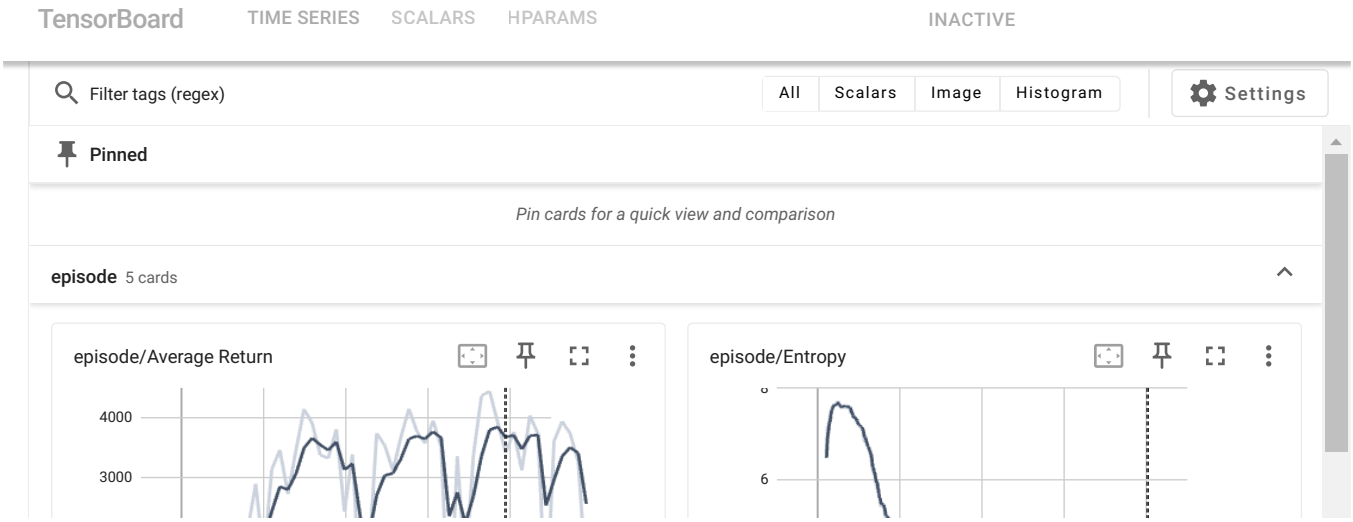
```
class PPO(LightningModule):

  def __init__(self, env_name, num_envs=2048, episode_length=1_000,
               batch_size=1024, hidden_size=256, samples_per_epoch=5,
               policy_lr=1e-4, value_lr=1e-3, epoch_repeat=8, epsilon=0.3,
               gamma=0.99, lamb=0.95, entropy_coef=0.2, optim=AdamW):

    super().__init__()

    self.env = create_env(env_name, num_envs=num_envs, episode_length=episode_length)
    test_env = gym.make(env_name, episode_length=episode_length)
    test_env = to_torch.JaxToTorchWrapper(test_env, device=device)
    self.test_env = NormalizeObservation(test_env)
    self.test_env.obs_rms = self.env.obs_rms

    obs_size = self.env.observation_space.shape[1]
    action_dims = self.env.action_space.shape[1]

    self.policy = GradientPolicy(obs_size, action_dims, hidden_size)
    self.value_net = ValueNet(obs_size, hidden_size)
    self.target_value_net = copy.deepcopy(self.value_net)

    self.dataset = RLDataset(self.env, self.policy, self.target_value_net,
                             samples_per_epoch, gamma, lamb, epoch_repeat)

    self.automatic_optimization = False

    self.save_hyperparameters()
    self.videos = []

  def configure_optimizers(self):
    value_opt = self.hparams.optim(self.value_net.parameters(), lr=self.hparams.value_lr)
    policy_opt = self.hparams.optim(self.policy.parameters(), lr=self.hparams.policy_lr)
    return value_opt, policy_opt

  def train_dataloader(self):
    return DataLoader(dataset=self.dataset, batch_size=self.hparams.batch_size)

  # Training step.
  def training_step(self, batch, batch_idx):
    obs_b, loc_b, scale_b, action_b, reward_b, gae_b, target_b = batch
```

```
    # Access optimizers
    opt_value, opt_policy = self.optimizers()

    # Value network optimization
    state_values = self.value_net(obs_b)
    loss_value = F.smooth_l1_loss(state_values, target_b)

    # Manual optimization for value network
    opt_value.zero_grad()
    self.manual_backward(loss_value)
    opt_value.step()

   # Policy network optimization
    new_loc, new_scale = self.policy(obs_b)
    dist = Normal(new_loc, new_scale)
    log_prob = dist.log_prob(action_b).sum(dim=-1, keepdim=True)

    prev_dist = Normal(loc_b, scale_b)
    prev_log_prob = prev_dist.log_prob(action_b).sum(dim=-1, keepdim=True)

    rho_s = torch.exp(log_prob - prev_log_prob)

    surrogate_1 = rho_s * gae_b
    surrogate_2 = rho_s.clip(1 - self.hparams.epsilon, 1 + self.hparams.epsilon) * gae_b
    policy_loss = - torch.minimum(surrogate_1, surrogate_2)

    entropy = dist.entropy().sum(dim=-1, keepdim=True)
    loss = policy_loss - self.hparams.entropy_coef * entropy

    # Manual optimization for policy network
    opt_policy.zero_grad()
    self.manual_backward(loss.mean())
    opt_policy.step()

    self.log("episode/Value Loss", loss_value)
    self.log("episode/Policy Loss", policy_loss.mean())
    self.log("episode/Entropy", entropy.mean())
    self.log("episode/Reward", reward_b.mean())

  def on_train_epoch_end(self):
    self.target_value_net.load_state_dict(self.value_net.state_dict())

    if self.current_epoch % 10 == 0:
      average_return = test_agent(self.test_env, self.hparams.episode_length, self.policy, episodes=1)
      self.log("episode/Average Return", average_return)

    if self.current_epoch % 50 == 0:
      video = create_video(self.test_env, self.hparams.episode_length, policy=self.policy)
      self.videos.append(video)
```

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/videos/': No such file or directory
The tensorboard extension is already loaded. To reload it, use:
  %reload_ext tensorboard
```




## Train the policy

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

```
algo = PPO('brax-halfcheetah-v0')

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=500,
    log_every_n_steps=1
)

trainer.fit(algo)
```

```
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
```