In [1]:

```
# Import necessary modules
import boto3
import sagemaker
import os
```

In [2]:

```
from sagemaker import get_execution_role
session =sagemaker.Session()
# store the current SageMaker session
# get IAM role
role=get_execution_role()
print(role)

bucket_name=session.default_bucket()
prefix = 'cancer-class'
```

```
arn:aws:iam::172268057478:role/service-role/AmazonSageMaker-ExecutionRole-202
10122T150167
```

In [3]:

```
# set prefix, a descriptive name for a directory for our train test data
data_dir = '../Capstone Project/data/'
prefix = 'cancer-class'
# upload all data to S3
test_location = session.upload_data(os.path.join(data_dir,'test.csv'),key_prefix=prefix)
train_location= session.upload_data(os.path.join(data_dir,'train.csv'),key_prefix=prefix)
```

In [13]:

```
# Create an estimtor
# your import and estimator code, here

from sagemaker.pytorch import PyTorch
output_path = 's3://{}/{}'.format(bucket_name, prefix)
estimator = PyTorch(entry_point="train.py",
                    source_dir="source_pytorch",
                    role=role,
                    framework_version='1.0',
                    py_version = 'py3',
                    sagemaker_session = session,
                    output_path = output_path,
                    train_instance_count=1,
                    train_instance_type='ml.p2.xlarge',
                    hyperparameters= {'input_features':5,
                                      'hidden_dim':5, #change this to 6
                                      'output_dim':1,
                                      'epochs':160
                                     }
                   )
```

```
train_instance_count has been renamed in sagemaker>=2.
See: https://sagemaker.readthedocs.io/en/stable/v2.html for details.
train_instance_type has been renamed in sagemaker>=2.
See: https://sagemaker.readthedocs.io/en/stable/v2.html for details.
```

In [14]:

```
%time

# Train your estimator on S3 training data
s3_input_train = sagemaker.TrainingInput(s3_data = train_location, content_type='csv')
estimator.fit({'train':s3_input_train})
```

```
CPU times: user 3 µs, sys: 0 ns, total: 3 µs
Wall time: 6.91 µs
2021-03-15 00:24:15 Starting - Starting the training job...
2021-03-15 00:24:40 Starting - Launching requested ML instancesProfilerReport
-1615767855: InProgress
......
2021-03-15 00:25:41 Starting - Preparing the instances for trainin
g.............
2021-03-15 00:27:42 Downloading - Downloading input data
2021-03-15 00:27:42 Training - Downloading the training image.....bash: canno
t set terminal process group (-1): Inappropriate ioctl for device
bash: no job control in this shell
2021-03-15 00:28:30,058 sagemaker-containers INFO     Imported framework sage
maker_pytorch_container.training
2021-03-15 00:28:30,084 sagemaker_pytorch_container.training INFO     Block u
ntil all host DNS lookups succeed.

2021-03-15 00:28:43 Training - Training image download completed. Training in
progress.2021-03-15 00:28:33,117 sagemaker_pytorch_container.training INFO
Invoking user training script.
2021-03-15 00:28:33,480 sagemaker-containers INFO     Module train does not p
rovide a setup.py.
Generating setup.py
2021-03-15 00:28:33,481 sagemaker-containers INFO     Generating setup.cfg
2021-03-15 00:28:33,481 sagemaker-containers INFO     Generating MANIFEST.in
2021-03-15 00:28:33,481 sagemaker-containers INFO     Installing module with
 the following command:
/usr/bin/python -m pip install -U .
Processing /opt/ml/code
Building wheels for collected packages: train
  Running setup.py bdist_wheel for train: started
  Running setup.py bdist_wheel for train: finished with status 'done'
  Stored in directory: /tmp/pip-ephem-wheel-cache-x_h05p5e/wheels/35/24/16/37
574d11bf9bde50616c67372a334f94fa8356bc7164af8ca3
Successfully built train
Installing collected packages: train
Successfully installed train-1.0.0
You are using pip version 18.1, however version 21.0.1 is available.
You should consider upgrading via the 'pip install --upgrade pip' command.
2021-03-15 00:28:35,578 sagemaker-containers INFO     Invoking user script

Training Env:

{
    "additional_framework_parameters": {},
    "channel_input_dirs": {
        "train": "/opt/ml/input/data/train"
    },
    "current_host": "algo-1",
    "framework_module": "sagemaker_pytorch_container.training:main",
    "hosts": [
        "algo-1"
    ],
    "hyperparameters": {
        "hidden_dim": 5,
        "input_features": 5,
        "epochs": 160,
```

```
        "output_dim": 1
    },
    "input_config_dir": "/opt/ml/input/config",
    "input_data_config": {
        "train": {
            "ContentType": "csv",
            "TrainingInputMode": "File",
            "S3DistributionType": "FullyReplicated",
            "RecordWrapperType": "None"
        }
    },
    "input_dir": "/opt/ml/input",
    "is_master": true,
    "job_name": "sagemaker-pytorch-2021-03-15-00-24-15-501",
    "log_level": 20,
    "master_hostname": "algo-1",
    "model_dir": "/opt/ml/model",
    "module_dir": "s3://sagemaker-us-east-1-172268057478/sagemaker-pytorch-20
21-03-15-00-24-15-501/source/sourcedir.tar.gz",
    "module_name": "train",
    "network_interface_name": "eth0",
    "num_cpus": 4,
    "num_gpus": 1,
    "output_data_dir": "/opt/ml/output/data",
    "output_dir": "/opt/ml/output",
    "output_intermediate_dir": "/opt/ml/output/intermediate",
    "resource_config": {
        "current_host": "algo-1",
        "hosts": [
            "algo-1"
        ],
        "network_interface_name": "eth0"
    },
    "user_entry_point": "train.py"
}

Environment variables:

SM_HOSTS=["algo-1"]
SM_NETWORK_INTERFACE_NAME=eth0
SM_HPS={"epochs":160,"hidden_dim":5,"input_features":5,"output_dim":1}
SM_USER_ENTRY_POINT=train.py
SM_FRAMEWORK_PARAMS={}
SM_RESOURCE_CONFIG={"current_host":"algo-1","hosts":["algo-1"],"network_inter
face_name":"eth0"}
SM_INPUT_DATA_CONFIG={"train":{"ContentType":"csv","RecordWrapperType":"Non
e","S3DistributionType":"FullyReplicated","TrainingInputMode":"File"}}
SM_OUTPUT_DATA_DIR=/opt/ml/output/data
SM_CHANNELS=["train"]
SM_CURRENT_HOST=algo-1
SM_MODULE_NAME=train
SM_LOG_LEVEL=20
SM_FRAMEWORK_MODULE=sagemaker_pytorch_container.training:main
SM_INPUT_DIR=/opt/ml/input
SM_INPUT_CONFIG_DIR=/opt/ml/input/config
SM_OUTPUT_DIR=/opt/ml/output
SM_NUM_CPUS=4
```

```
SM_NUM_GPUS=1
SM_MODEL_DIR=/opt/ml/model
SM_MODULE_DIR=s3://sagemaker-us-east-1-172268057478/sagemaker-pytorch-2021-03
-15-00-24-15-501/source/sourcedir.tar.gz
SM_TRAINING_ENV={"additional_framework_parameters":{},"channel_input_dirs":
{"train":"/opt/ml/input/data/train"},"current_host":"algo-1","framework_modul
e":"sagemaker_pytorch_container.training:main","hosts":["algo-1"],"hyperparam
eters":{"epochs":160,"hidden_dim":5,"input_features":5,"output_dim":1},"input
_config_dir":"/opt/ml/input/config","input_data_config":{"train":{"ContentTyp
e":"csv","RecordWrapperType":"None","S3DistributionType":"FullyReplicated","T
rainingInputMode":"File"}},"input_dir":"/opt/ml/input","is_master":true,"job_
name":"sagemaker-pytorch-2021-03-15-00-24-15-501","log_level":20,"master_host
name":"algo-1","model_dir":"/opt/ml/model","module_dir":"s3://sagemaker-us-ea
st-1-172268057478/sagemaker-pytorch-2021-03-15-00-24-15-501/source/sourcedir.
tar.gz","module_name":"train","network_interface_name":"eth0","num_cpus":4,"n
um_gpus":1,"output_data_dir":"/opt/ml/output/data","output_dir":"/opt/ml/outp
ut","output_intermediate_dir":"/opt/ml/output/intermediate","resource_confi
g":{"current_host":"algo-1","hosts":["algo-1"],"network_interface_name":"eth
0"},"user_entry_point":"train.py"}
SM_USER_ARGS=["--epochs","160","--hidden_dim","5","--input_features","5","--o
utput_dim","1"]
SM_OUTPUT_INTERMEDIATE_DIR=/opt/ml/output/intermediate
SM_CHANNEL_TRAIN=/opt/ml/input/data/train
SM_HP_HIDDEN_DIM=5
SM_HP_INPUT_FEATURES=5
SM_HP_EPOCHS=160
SM_HP_OUTPUT_DIM=1
PYTHONPATH=/usr/local/bin:/usr/lib/python36.zip:/usr/lib/python3.6:/usr/lib/p
ython3.6/lib-dynload:/usr/local/lib/python3.6/dist-packages:/usr/lib/python3/
dist-packages

Invoking script with the following command:

/usr/bin/python -m train --epochs 160 --hidden_dim 5 --input_features 5 --out
put_dim 1


Using device cuda.
Get train data loader.
Epoch: 1, Loss: 0.4374334741383791
Epoch: 2, Loss: 0.3048462275415659
Epoch: 3, Loss: 0.26107760053128004
Epoch: 4, Loss: 0.23207261729985476
Epoch: 5, Loss: 0.2064603335224092
Epoch: 6, Loss: 0.2010144186206162
Epoch: 7, Loss: 0.1917874976992607
Epoch: 8, Loss: 0.1841118335723877
Epoch: 9, Loss: 0.17689115200191735
Epoch: 10, Loss: 0.16207092879340051
Epoch: 11, Loss: 0.16850282535888256
Epoch: 12, Loss: 0.17888695681467653
Epoch: 13, Loss: 0.13538551330566406
Epoch: 14, Loss: 0.1622675877995789
Epoch: 15, Loss: 0.12792302700690925
Epoch: 16, Loss: 0.13055760969873517
Epoch: 17, Loss: 0.15648520134855062
Epoch: 18, Loss: 0.19517087067943067
```

```
Epoch: 19, Loss: 0.14933817808050662
Epoch: 20, Loss: 0.1317520899232477
Epoch: 21, Loss: 0.12612339458428323
Epoch: 22, Loss: 0.12328000294510275
Epoch: 23, Loss: 0.11412884495221079
Epoch: 24, Loss: 0.12724476722069084
Epoch: 25, Loss: 0.12025224715471268
Epoch: 26, Loss: 0.12677132759708912
Epoch: 27, Loss: 0.13087865854613484
Epoch: 28, Loss: 0.11553791556507349
Epoch: 29, Loss: 0.12928105446044355
Epoch: 30, Loss: 0.12065780314151198
Epoch: 31, Loss: 0.12254073361400515
Epoch: 32, Loss: 0.11392505186377093
Epoch: 33, Loss: 0.11542091713054106
Epoch: 34, Loss: 0.10339704948710278
Epoch: 35, Loss: 0.13215526471612976
Epoch: 36, Loss: 0.12878501329105346
Epoch: 37, Loss: 0.10927909564925357
Epoch: 38, Loss: 0.11106782534625381
Epoch: 39, Loss: 0.11911328566493466
Epoch: 40, Loss: 0.11365296215517447
Epoch: 41, Loss: 0.11159275927348063
Epoch: 42, Loss: 0.11701437142910436
Epoch: 43, Loss: 0.11740770098986104
Epoch: 44, Loss: 0.10715342247858643
Epoch: 45, Loss: 0.11257903852965682
Epoch: 46, Loss: 0.11894611421739683
Epoch: 47, Loss: 0.12451589796692134
Epoch: 48, Loss: 0.11142903035506606
Epoch: 49, Loss: 0.11900548527482897
Epoch: 50, Loss: 0.12312367042759434
Epoch: 51, Loss: 0.12363722575828433
Epoch: 52, Loss: 0.12143976431107148
Epoch: 53, Loss: 0.10602222438901662
Epoch: 54, Loss: 0.1060070700244978
Epoch: 55, Loss: 0.10296947414753958
Epoch: 56, Loss: 0.11265300472732634
Epoch: 57, Loss: 0.11474736959207803
Epoch: 58, Loss: 0.12976451156428084
Epoch: 59, Loss: 0.10631646403344348
Epoch: 60, Loss: 0.10101622922811657
Epoch: 61, Loss: 0.11085960768396035
Epoch: 62, Loss: 0.10504162797005848
Epoch: 63, Loss: 0.10257098004221916
Epoch: 64, Loss: 0.11504737323848531
Epoch: 65, Loss: 0.11462046990636736
Epoch: 66, Loss: 0.1155908887158148
Epoch: 67, Loss: 0.10547262327745557
Epoch: 68, Loss: 0.1074382095132023
Epoch: 69, Loss: 0.09895823979750276
Epoch: 70, Loss: 0.1069446183857508
Epoch: 71, Loss: 0.10429856982082128
Epoch: 72, Loss: 0.15483179299626498
Epoch: 73, Loss: 0.09958587239962072
Epoch: 74, Loss: 0.09693259972846135
Epoch: 75, Loss: 0.10173361813649535
```

```
Epoch: 76, Loss: 0.09250547463307157
Epoch: 77, Loss: 0.0904126379173249
Epoch: 78, Loss: 0.08729839193401859
Epoch: 79, Loss: 0.09378168497933075
Epoch: 80, Loss: 0.09312541614053771
Epoch: 81, Loss: 0.14016235418384895
Epoch: 82, Loss: 0.09366122932406143
Epoch: 83, Loss: 0.08767776273889467
Epoch: 84, Loss: 0.08770668413490057
Epoch: 85, Loss: 0.08851073577534407
Epoch: 86, Loss: 0.09014127082191407
Epoch: 87, Loss: 0.09282814165344462
Epoch: 88, Loss: 0.09735903689288535
Epoch: 89, Loss: 0.08716438188566826
Epoch: 90, Loss: 0.08668998202192597
Epoch: 91, Loss: 0.08488278172444552
Epoch: 92, Loss: 0.08471327542793006
Epoch: 93, Loss: 0.08429163818364031
Epoch: 94, Loss: 0.08394273811718449
Epoch: 95, Loss: 0.09063559700734913
Epoch: 96, Loss: 0.08834380460320972
Epoch: 97, Loss: 0.08577191396616399
Epoch: 98, Loss: 0.09559030301170424
Epoch: 99, Loss: 0.1558416201500222
Epoch: 100, Loss: 0.11980582819669508
Epoch: 101, Loss: 0.10004724410828203
Epoch: 102, Loss: 0.09019713762099854
Epoch: 103, Loss: 0.12321694145794027
Epoch: 104, Loss: 0.11128187178983354
Epoch: 105, Loss: 0.11140104451915249
Epoch: 106, Loss: 0.1336655988590792
Epoch: 107, Loss: 0.0957030423800461
Epoch: 108, Loss: 0.11680371996480972
Epoch: 109, Loss: 0.10558766819303855
Epoch: 110, Loss: 0.10203970570582896
Epoch: 111, Loss: 0.1011468727258034
Epoch: 112, Loss: 0.10071751200594008
Epoch: 113, Loss: 0.10414272563066333
Epoch: 114, Loss: 0.10940568157238886
Epoch: 115, Loss: 0.10055784540018067
Epoch: 116, Loss: 0.10331875429255888
Epoch: 117, Loss: 0.10509718696121126
Epoch: 118, Loss: 0.1001111235236749
Epoch: 119, Loss: 0.1067656148225069
Epoch: 120, Loss: 0.12036150504136459
Epoch: 121, Loss: 0.10758436331525445
Epoch: 122, Loss: 0.15268804654479026
Epoch: 123, Loss: 0.12109345116186887
Epoch: 124, Loss: 0.11281123532680795
Epoch: 125, Loss: 0.10120895334985107
Epoch: 126, Loss: 0.10231580300023779
Epoch: 127, Loss: 0.10896592367207632
Epoch: 128, Loss: 0.10593189254868776
Epoch: 129, Loss: 0.10005880816606805
Epoch: 130, Loss: 0.10632971213199198
Epoch: 131, Loss: 0.15948718631407247
Epoch: 132, Loss: 0.14675211724825205
```

```
Epoch: 133, Loss: 0.13493956507882104
Epoch: 134, Loss: 0.11290856221457943
Epoch: 135, Loss: 0.10749372597783804
Epoch: 136, Loss: 0.11820736583322286
Epoch: 137, Loss: 0.12513974676840006
Epoch: 138, Loss: 0.10050144074484706
Epoch: 139, Loss: 0.10123830561060458
Epoch: 140, Loss: 0.10317098218947648
Epoch: 141, Loss: 0.09709527840605006
Epoch: 142, Loss: 0.0911611701361835
Epoch: 143, Loss: 0.09818543614819646
Epoch: 144, Loss: 0.09421929444652051
Epoch: 145, Loss: 0.09597092896001413
Epoch: 146, Loss: 0.09974960676627234
Epoch: 147, Loss: 0.0885001233080402
Epoch: 148, Loss: 0.10118985936278477
Epoch: 149, Loss: 0.09362403788836673
Epoch: 150, Loss: 0.09057853352278471
Epoch: 151, Loss: 0.09815730473492294
Epoch: 152, Loss: 0.11813456577947364
Epoch: 153, Loss: 0.0954888632928487
Epoch: 154, Loss: 0.09508721776073799
Epoch: 155, Loss: 0.09350353562040255
Epoch: 156, Loss: 0.09221316847251729
Epoch: 157, Loss: 0.09408691357239149
Epoch: 158, Loss: 0.0958908787346445
Epoch: 159, Loss: 0.09016155251883902
Epoch: 160, Loss: 0.08222302198410034
2021-03-15 00:28:56,780 sagemaker-containers INFO     Reporting training SUCC
ESS

2021-03-15 00:29:07 Uploading - Uploading generated training model
2021-03-15 00:29:07 Completed - Training job completed
Training seconds: 108
Billable seconds: 108
```

In [ ]:

```
# deploy your model to create a predictor
%time
predictor = estimator.deploy(initial_instance_count=1,instance_type='ml.m4.xlarge')
```

In [15]:

```
import os
import pandas as pd

# read in test data, assuming it is stored locally
test_data = pd.read_csv(os.path.join(data_dir, "test.csv"), header=None, names=None)
# labels are in the first column
test_y = test_data.iloc[:,0]
test_x = test_data.iloc[:,1:]
```

In [16]:

```
from sklearn.metrics import accuracy_score
import numpy as np
test_y_preds = predictor.predict(test_x.values.astype(np.float32))
test_y_preds = [round(num) for num in test_y_preds.squeeze()]
```

In [17]:

```
# debug for overfitting
import numpy as np
train_data = pd.read_csv(os.path.join(data_dir, "train.csv"), header=None, names=None)
train_y = train_data.iloc[:,0]
train_x = train_data.iloc[:,1:]
train_y_preds = predictor.predict(train_x.values.astype(np.float32))
train_y_preds = [round(num) for num in train_y_preds.squeeze()]
accuracy = accuracy_score(train_y, train_y_preds)
print("%4.2f" % (100*accuracy), "%" )
```

```
96.98 %
```

In [18]:

```
# Second: calculate the test accuracy
accuracy = accuracy_score(test_y, test_y_preds)

print("%4.2f %%" % (100* accuracy) )


## print out the array of predicted and true labels, if you want
'''
print('\nPredicted class labels: ')
print(test_y_preds)
print('\nTrue class labels: ')

print(test_y.values)
'''
```

```
97.66 %
```

Out[18]:

```
"\nprint('\nPredicted class labels: ')\nprint(test_y_preds)\nprint('\nTrue cl
ass labels: ')\n\nprint(test_y.values)\n"
```

In [ ]:

```
```