Lab 1. PyTorch and ANNs

This lab is a warm up to get you used to the PyTorch programming environment used in the course, and also to help you review and renew your knowledge of Python and relevant Python libraries. The lab must be done individually. Please recall that the University of Toronto plagarism rules apply.

By the end of this lab, you should be able to:

1. Be able to perform basic PyTorch tensor operations.
2. Be able to load data into PyTorch
3. Be able to configure an Artificial Neural Network (ANN) using PyTorch
4. Be able to train ANNs using PyTorch
5. Be able to evaluate different ANN configuations

You will need to use numpy and PyTorch documentations for this assignment:

- https://docs.scipy.org/doc/numpy/reference/ (https://docs.scipy.org/doc/numpy/reference/)
- https://pytorch.org/docs/stable/torch.html (https://pytorch.org/docs/stable/torch.html)

You can also reference Python API documentations freely.

What to submit ¶

Submit a PDF file containing all your code, outputs, and write-up from parts 1-5. You can produce a PDF of your Google Colab file by going to `File -> Print` and then save as PDF. The Colab instructions has more information.

Do not submit any other files produced by your code.

Include a link to your colab file in your submission.

Please use Google Colab to complete this assignment. If you want to use Jupyter Notebook, please complete the assignment and upload your Jupyter Notebook file to Google Colab for submission.

Adjust the scaling to ensure that the text is not cutoff at the margins.

Colab Link

Submit make sure to include a link to your colab file here

Colab Link: https://colab.research.google.com/drive/1WikemeOUpir4xw0TZ67SkjLXOq8lbCI9?usp=sharing (https://colab.research.google.com/drive/1WikemeOUpir4xw0TZ67SkjLXOq8lbCI9?usp=sharing)

Part 1. Python Basics [3 pt]

The purpose of this section is to get you used to the basics of Python, including working with functions, numbers, lists, and strings.

Note that we **will** be checking your code for clarity and efficiency.

If you have trouble with this part of the assignment, please review http://cs231n.github.io/python-numpy-tutorial/ (http://cs231n.github.io/python-numpy-tutorial/)

Part (a) -- 1pt

Write a function `sum_of_cubes` that computes the sum of cubes up to `n` . If the input to `sum_of_cubes` invalid (e.g. negative or non-integer `n`), the function should print out `"Invalid input"` and return `-1` .

In [6]:
```
def sum_of_cubes(n):
    """Return the sum (1^3 + 2^3 + 3^3 + ... + n^3)

    Precondition: n > 0, type(n) == int

    >>> sum_of_cubes(3)
    36
    >>> sum_of_cubes(1)
    1
    """

    if type(n) != int or n <= 0:
      print("Invalid input")
      return -1

    sum = 0
    for i in range(1, n+1):
      sum += i**3
    return sum
```

Part (b) -- 1pt

Write a function `word_lengths` that takes a sentence (string), computes the length of each word in that sentence, and returns the length of each word in a list. You can assume that words are always separated by a space character `" "` .

Hint: recall the `str.split` function in Python. If you arenot sure how this function works, try typing `help(str.split)` into a Python shell, or check out https://docs.python.org/3.6/library/stdtypes.html#str.split (https://docs.python.org/3.6/library/stdtypes.html#str.split)

In [ ]:
```
help(str.split)
```

```
Help on method_descriptor:

split(self, /, sep=None, maxsplit=-1)
    Return a list of the words in the string, using sep as the delimiter stri
ng.

    sep
      The delimiter according which to split the string.
      None (the default value) means split according to any whitespace,
      and discard empty strings from the result.
    maxsplit
      Maximum number of splits to do.
      -1 (the default value) means no limit.
```

In [8]:
```
def word_lengths(sentence):
    """Return a list containing the length of each word in
    sentence.

    >>> word_lengths("welcome to APS360!")
    [7, 2, 7]
    >>> word_lengths("machine learning is so cool")
    [7, 8, 2, 2, 4]
    """

    lengths = []
    for i in sentence.split():
      lengths.append(len(i))

    return lengths
```

Part (c) -- 1pt

Write a function `all_same_length` that takes a sentence (string), and checks whether every word in the string is the same length. You should call the function `word_lengths` in the body of this new function.

In [14]:
```
def all_same_length(sentence):
    """Return True if every word in sentence has the same
    length, and False otherwise.

    >>> all_same_length("all same length")
    False
    >>> word_lengths("hello world")
    True
    """

    lengths = word_lengths(sentence)
    for i in range(1, len(lengths)):
      if lengths[i] != lengths[0]:
        return False
    return True
```

Out[14]:
```
False
```

Part 2. NumPy Exercises [5 pt]

In this part of the assignment, you'll be manipulating arrays usign NumPy. Normally, we use the shorter name `np` to represent the package `numpy` .

In [15]:
```
import numpy as np
```

Part (a) -- 1pt

The below variables `matrix` and `vector` are numpy arrays. Explain what you think `<NumpyArray>.size` and `<NumpyArray>.shape` represent.

In [17]:
```
matrix = np.array([[1., 2., 3., 0.5],
                   [4., 5., 0., 0.],
                   [-1., -2., 1., 1.]])
vector = np.array([2., 0., 1., -2.])
```

In [ ]:
```
matrix.size
```

Out[ ]:
```
12
```

In [ ]:
```
matrix.shape
```

Out[ ]:
```
(3, 4)
```

In [ ]:
```
vector.size
```

Out[ ]:
```
4
```

In [ ]:
```
vector.shape
```

Out[ ]:
```
(4,)
```

<NumpyArray>.size returns the number of elements in the numpy array as an int.

<NumpyArray>.shape returns the shape of the array (length of the array in each dimension) as a tuple.

Part (b) -- 1pt

Perform matrix multiplication `output = matrix x vector` by using for loops to iterate through the columns and rows. Do not use any builtin NumPy functions. Cast your output into a NumPy array, if it isn't one already.

Hint: be mindful of the dimension of output

In [18]:
```
rows, columns = matrix.shape
output = []

for i in range(rows):
  sum = 0
  for j in range(columns):
    sum += matrix[i, j] * vector[j]
  output.append(sum)

output = np.asarray(output)
```

In [19]:
```
output
```

Out[19]:
```
array([ 4.,  8., -3.])
```

Part (c) -- 1pt

Perform matrix multiplication `output2 = matrix x vector` by using the function `numpy.dot` .

We will never actually write code as in part(c), not only because `numpy.dot` is more concise and easier to read/write, but also performance-wise `numpy.dot` is much faster (it is written in C and highly optimized). In general, we will avoid for loops in our code.

In [20]:
```
output2 = np.dot(matrix, vector)
```

In [21]:
```
output2
```

Out[21]:
```
array([ 4.,  8., -3.])
```

Part (d) -- 1pt

As a way to test for consistency, show that the two outputs match.

In [22]:
```
if np.array_equal(output, output2):
  print("The two outputs are the same")
```

```
The two outputs are the same
```

Part (e) -- 1pt

Show that using `np.dot` is faster than using your code from part (c).

You may find the below code snippit helpful:

In [ ]:
```
import time

# record the time before running code
start_time = time.time()

# place code to run here
for i in range(10000):
    99*99

# record the time after the code is run
end_time = time.time()

# compute the difference
diff = end_time - start_time
diff
```

Out[ ]:
```
0.0009815692901611328
```

In [ ]:
```
# Loop method
start_time_loop = time.time()
rows, columns = matrix.shape
output = []

for i in range(rows):
  sum = 0
  for j in range(columns):
    sum += matrix[i, j] * vector[j]
  output.append(sum)

output = np.asarray(output)
end_time_loop = time.time()

# np.dot method
start_time_dot = time.time()
output2 = np.dot(matrix, vector)
end_time_dot = time.time()

# Print results
print(f"The time using loops is: {end_time_loop-start_time_loop}")
print(f"The time using np.dot is: {end_time_dot-start_time_dot}")

# np.dot is generally faster, but sometimes the loop will be faster
# the inconsistency may be due to the small size of the matrix and vector
```

```
The time using loops is: 0.00026440620422363281
The time using np.dot is: 7.915496826171875e-05
```

Part 3. Images [6 pt]

A picture or image can be represented as a NumPy array of “pixels”, with dimensions H × W × C, where H is the height of the image, W is the width of the image, and C is the number of colour channels. Typically we will use an image with channels that give the the Red, Green, and Blue “level” of each pixel, which is referred to with the short form RGB.

You will write Python code to load an image, and perform several array manipulations to the image and visualize their effects.

In [24]:
```
import matplotlib.pyplot as plt
```

Part (a) -- 1 pt

This is a photograph of a dog whose name is Mochi.

Load the image from its url (https://drive.google.com/uc?export=view&id=1oaLVR2hr1_qzpKQ47i9rVUIklwbDcews (https://drive.google.com/uc?export=view&id=1oaLVR2hr1_qzpKQ47i9rVUIklwbDcews)) into the variable `img` using the `plt.imread` function.

Hint: You can enter the URL directly into the `plt.imread` function as a Python string.

In [26]:
```
img = plt.imread("https://drive.google.com/uc?export=view&id=1oaLVR2hr1_qzpKQ47i9rVUIklwbDcews")
```

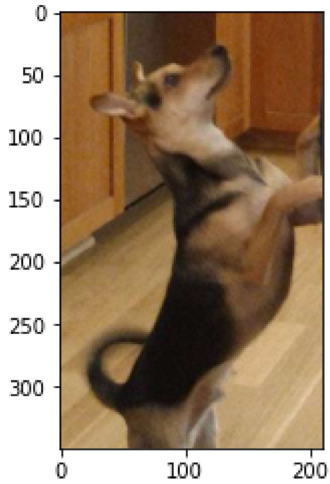
Part (b) -- 1pt

Use the function `plt.imshow` to visualize `img` .

This function will also show the coordinate system used to identify pixels. The origin is at the top left corner, and the first dimension indicates the Y (row) direction, and the second dimension indicates the X (column) dimension.

In [27]:
```
plt.imshow(img)
```

Out[27]: <matplotlib.image.AxesImage at 0x7fad56571f70>

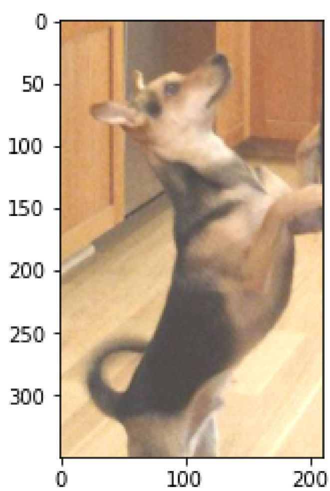


Part (c) -- 2pt

Modify the image by adding a constant value of 0.25 to each pixel in the `img` and store the result in the variable `img_add`. Note that, since the range for the pixels needs to be between [0, 1], you will also need to clip img_add to be in the range [0, 1] using `numpy.clip`. Clipping sets any value that is outside of the desired range to the closest endpoint. Display the image using `plt.imshow`.

In [28]:
```
img_add = np.clip(img + 0.25, 0, 1)
plt.imshow(img_add)
```

Out[28]: <matplotlib.image.AxesImage at 0x7fad554947f0>



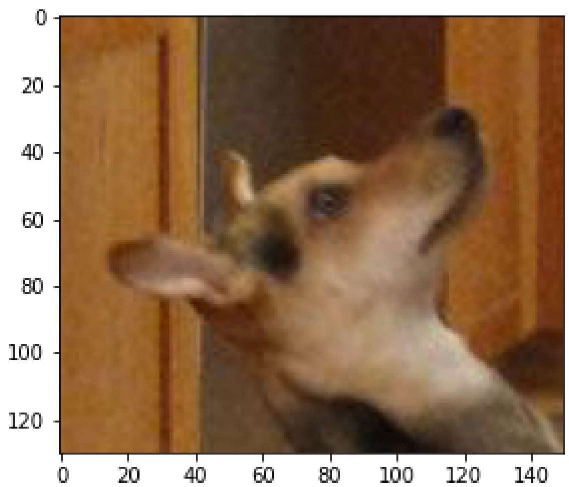
Part (d) -- 2pt

Crop the **original** image (`img` variable) to a 130 x 150 image including Mochi's face. Discard the alpha colour channel (i.e. resulting `img_cropped` should **only have RGB channels**)

Display the image.

In [36]:
```
img_cropped = img[0:130, 10:160, 0:3]
plt.imshow(img_cropped)
```

Out[36]: `<matplotlib.image.AxesImage at 0x7fad5523d2b0>`



Part 4. Basics of PyTorch [6 pt]

PyTorch is a Python-based neural networks package. Along with tensorflow, PyTorch is currently one of the most popular machine learning libraries.

PyTorch, at its core, is similar to Numpy in a sense that they both try to make it easier to write codes for scientific computing achieve improved performance over vanilla Python by leveraging highly optimized C back-end. However, compare to Numpy, PyTorch offers much better GPU support and provides many high-level features for machine learning. Technically, Numpy can be used to perform almost every thing PyTorch does. However, Numpy would be a lot slower than PyTorch, especially with CUDA GPU, and it would take more effort to write machine learning related code compared to using PyTorch.

In [37]:
```
import torch
```

Part (a) -- 1 pt

Use the function `torch.from_numpy` to convert the numpy array `img_cropped` into a PyTorch tensor. Save the result in a variable called `img_torch` .

In [38]:
```
img_torch = torch.from_numpy(img_cropped)
```

Part (b) -- 1pt

Use the method `<Tensor>.shape` to find the shape (dimension and size) of `img_torch` .

In [39]:
```
img_torch.shape #130x150x3
```

Out[39]:
```
torch.Size([130, 150, 3])
```

Part (c) -- 1pt

How many floating-point numbers are stored in the tensor `img_torch` ?

In [40]:
```
# torch tensor only contains elements of a single data type
img_torch.numel() # 58500 floating point numbers
```

Out[40]:
```
58500
```

Part (d) -- 1 pt

What does the code `img_torch.transpose(0,2)` do? What does the expression return? Is the original variable `img_torch` updated? Explain.

In [ ]:
```
# The code will return a new tensor with the values in dimensions 0 and 2 swap
ped with each other.
# Since the code returns a new tensor, the original variable img_torch is not
changed by the transpose.
```

Part (e) -- 1 pt

What does the code `img_torch.unsqueeze(0)` do? What does the expression return? Is the original variable `img_torch` updated? Explain.

In [ ]:
```
# The code will return a new tensor with a dimension of size 1 inserted at axi
s 0 (before the other dimensions).
# Since the code returns a new tensor, the original variable img_torch is not
changed.
```

Part (f) -- 1 pt

Find the maximum value of `img_torch` along each colour channel? Your output should be a one-dimensional PyTorch tensor with exactly three values.

Hint: lookup the function `torch.max`.

In [ ]:
```
torch.max(torch.max(img_torch, 1)[0], 1)[0]
```

Out[ ]:
```
tensor([0.8941, 0.7882, 0.6745])
```

Part 5. Training an ANN [10 pt]

The sample code provided below is a 2-layer ANN trained on the MNIST dataset to identify digits less than 3 or greater than and equal to 3. Modify the code by changing any of the following and observe how the accuracy and error are affected:

- number of training iterations
- number of hidden units
- numbers of layers
- types of activation functions
- learning rate

Please select at least three different options from the list above. For each option, please select two to three different parameters and provide a table.

In [ ]:
```
import torch
import torch.nn as nn
import torch.nn.functional as F
from torchvision import datasets, transforms
import matplotlib.pyplot as plt # for plotting
import torch.optim as optim

torch.manual_seed(1) # set the random seed

# define a 2-layer artificial neural network
class Pigeon(nn.Module):
    def __init__(self):
        super(Pigeon, self).__init__()
        self.layer1 = nn.Linear(28 * 28, 30)
        self.layer2 = nn.Linear(30, 1)
    def forward(self, img):
        flattened = img.view(-1, 28 * 28)
        activation1 = self.layer1(flattened)
        activation1 = F.relu(activation1)
        activation2 = self.layer2(activation1)
        return activation2

pigeon = Pigeon()

# load the data
mnist_data = datasets.MNIST('data', train=True, download=True)
mnist_data = list(mnist_data)
mnist_train = mnist_data[:1000]
mnist_val   = mnist_data[1000:2000]
img_to_tensor = transforms.ToTensor()


# simplified training code to train `pigeon` on the "small digit recognition"
task
criterion = nn.BCEWithLogitsLoss()
optimizer = optim.SGD(pigeon.parameters(), lr=0.007, momentum=0.9)

num_iterations = 1 ### ISSUE
for i in range(num_iterations):
  for (image, label) in mnist_train:
      # actual ground truth: is the digit less than 3?
      actual = torch.tensor(label < 3).reshape([1,1]).type(torch.FloatTensor)
      # pigeon prediction
      out = pigeon(img_to_tensor(image)) # step 1-2
      # update the parameters based on the loss
      loss = criterion(out, actual)      # step 3
      loss.backward()                    # step 4 (compute the updates for eac
h parameter)
      optimizer.step()                   # step 4 (make the updates for each p
arameter)
      optimizer.zero_grad()              # a clean up step for PyTorch

# computing the error and accuracy on the training set
error = 0
for (image, label) in mnist_train:
    prob = torch.sigmoid(pigeon(img_to_tensor(image)))
```

```
    if (prob < 0.5 and label < 3) or (prob >= 0.5 and label >= 3):
        error += 1
print("Training Error Rate:", error/len(mnist_train))
print("Training Accuracy:", 1 - error/len(mnist_train))


# computing the error and accuracy on a test set
error = 0
for (image, label) in mnist_val:
    prob = torch.sigmoid(pigeon(img_to_tensor(image)))
    if (prob < 0.5 and label < 3) or (prob >= 0.5 and label >= 3):
        error += 1
print("Test Error Rate:", error/len(mnist_val))
print("Test Accuracy:", 1 - error/len(mnist_val))
```

```
Training Error Rate: 0.049
Training Accuracy: 0.951
Test Error Rate: 0.092
Test Accuracy: 0.908
```

Test Results

Hyperparameter	Setting	Training Accuracy/Error	Test Accuracy/Error
N/A	Default	96.4% / 3.6%	92.1% / 7.9%
# Training Iterations	5	98.9% / 1.1%	93.4% / 6.6%
	10	99.9% / 0.1%	94.1% / 5.9%
	15	99.9% / 0.1%	94.1% / 5.9%
# Hidden Units	60	96.9% / 3.1%	92% / 8%
	20	96.5% / 3.5%	91.2% / 8.8%
	40	96.9% / 3.1%	92.6% / 7.4%
Learning Rate	0.01	96.1% / 3.9%	91.8% / 8.2%
	0.002	95.6% / 4.4%	89.7% / 10.3%
	0.007	95.1% / 4.9%	90.8% / 9.2%

Part (a) -- 3 pt

Comment on which of the above changes resulted in the best accuracy on training data? What accuracy were you able to achieve?

Increasing the number of training iterations increased the training accuracy of the model. Running 15 iterations on the same data increased the accuracy from 96.4% (default settings) to 99.9% accuracy. This is not surprising since training the model repeatedly on the same data would greatly help it to remember corresponding labels.

Part (b) -- 3 pt

Comment on which of the above changes resulted in the best accuracy on testing data? What accuracy were you able to achieve?

Increasing the number of training iterations also increased the testing accuracy of the model. Running 15 iterations on the same training data increased the testing accuracy from 92.1% (at default settings) to 94.1%. Despite training the model for 15 iterations on the same data and the model producing a very high training accuracy, it seems to still be able to generalize well to unseen examples.

Part (c) -- 4 pt

Which model hyperparameters should you use, the ones from (a) or (b)?

Since running more training iterations improved both training and testing data, it is an easy decision to use at least 15 (or 10) training iterations. In this case, the benefit of reduced training error from running more iterations offsets the harm from an increased gap between the training and testing accuracies. However, running more than 15 iterations may produce significant overfitting that overcomes any benefit provided, harming the model's test accuracy.

If I had to choose between hyperparameters that result in a higher training accuracy or a higher test accuracy, I would choose the hyperparameters that produce a higher test accuracy. This is because when applying the model in real situations, it will generally be seeing new scenarios rather than examples from training, which is what the test data simulates. Thus, we want to optimize for testing accuracy so that the model can generalize better to unseen scenarios, performing better when it is actually deployed.