Chapter 5 Question

1. Why do we first resize to a large size on the CPU, and then to a smaller size on  
   the GPU?

First, we resize to large image on the CPU so that all image can have the same size and minimize data destruction and allow better further augmentation. After that, random crop ( to smaller size) is performed for augmentation. The image will look clearer

2. If you are not familiar with regular expressions, find a regular expression tutorial  
and some problem sets, and complete them. Have a look on the book’s website  
for suggestions.

Done  
3. What are the two ways in which data is most commonly provided for most deep  
learning datasets?

1. Individual files representing items of data, such as text documents or images, possibly organized into folders with filenames representing information about those items
2. Table of data ( such as CSV format), each row is an item, may include filename providing connection between data in the table and the data outside(image, text) of the table

4. Look up the documentation for L and try using a few of the new methods that it  
adds.

Done   
5. Look up the documentation for the Python pathlib module and try using a few  
methods of the Path class.

Done partially, not work with Collab  
6. Give two examples of ways that image transformations can degrade the quality of  
the data.

Example 1: Rotating an image by 45 degree can create empty zone on edge and will not teach the model anything

Example 2: Couple of augmentation method or interpolating pixels will reduce the quality

7. What method does fastai provide to view the data in a DataLoaders?

Dls.show\_batch(nrows = ?, ncols = ?)

8. What method does fastai provide to help you debug a DataBlock?

Datablocks(pets1).summary(path/”…”)

9. Should you hold off on training a model until you have thoroughly cleaned your  
data?

NO  
10. What are the two pieces that are combined into cross-entropy loss in PyTorch?

It combines log and softmax  
11. What are the two properties of activations that softmax ensures? Why is this  
important?

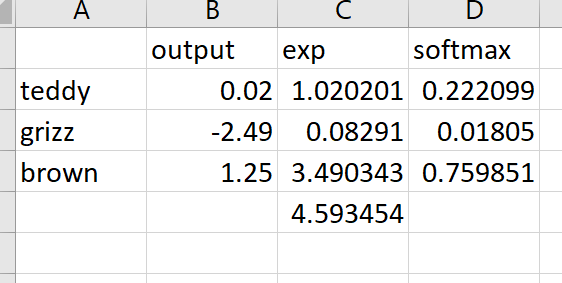
1. Softmax guarantee that probability of all categories add up to 1
2. Softmax guarantees that slightly larger, like 0.99 vs 0.999, it will magnify the difference.

If we don’t have this, the model wont care about 0.99 or 0.999

12. When might you want your activations to not have these two properties?

When you have multi-label classification problems (more than one label possible).

13. Calculate the exp and softmax columns of Figure 5-3 yourself (i.e., in a spread‐  
sheet, with a calculator, or in a notebook).



14. Why can’t we use torch.where to create a loss function for datasets where our  
label can have more than two categories?

Torch.where only return either this value or that value for two categories, so we can’t create loss function for data set with more than two categories.  
15. What is the value of log(–2)? Why?

Can’t, log only do positive value.

E^? cant be negative

16. What are two good rules of thumb for picking a learning rate from the learning  
rate finder?

The lowest point divide by 10

Last point where the lost clearly decreasing

17. What two steps does the fine\_tune method do?

Train the randomly added layer for one epoch, with other layers frozen

Unfreeze all the layers, and trains them for the number of epoch requested

18. In Jupyter Notebook, how do you get the source code for a method or function?  
doc()

19. What are discriminative learning rates?

Training different layer with different learning rates

20. How is a Python slice object interpreted when passed as a learning rate to  
fastai?  
The first value of the slice is the value that passed to the earliest layer in neural network

The second value will be for the last layer

The other layer will be the equidistant value between those two.  
21. Why is early stopping a poor choice when using 1cycle training?

Because later result could be better if you train it more.

22. What is the difference between resnet50 and resnet101?

There is 50 layers in resnet50

101 layer in resnet 101

23. What does to\_fp16 do?  
A technique to speed things up using mixed precision training, require less CPU so that it can speed up neural network training

Further Research  
1. Find the paper by Leslie Smith that introduced the learning rate finder, and read  
it.  
2. See if you can improve the accuracy of the classifier in this chapter. What’s the  
best accuracy you can achieve? Look on the forums and the book’s website to see  
what other students have achieved with this dataset and how they did it.