

5. What are the two ways in which data is most commonly provided, for most deep learning datasets?
Train and test are the dataset there and inside training dataset there are various folders which will contain the respected data which will be needed
6. Look up the documentation for L and try using a few of the new methods that it adds.
-----> L stands for the modified list in the fast ai it gives out the total number of elements at start and then the list of contained elements truncated at some point
7. Look up the documentation for the Python pathlib module and try using a few methods of the Path class.
8. Give two examples of ways that image transformations can degrade the quality of the data.
When we random crop the image maybe the original image get's cropped and the part of image which is unnecessary get stored in there sometimes when we do stretch the image data can be out of the order and it can be totally different from real world sample , the traditional cutting of the data can also be result in the image deformation / degradation which will be huge problem cause the imaeg pixels are not getting cut properly and some images get blurry
9. What method does fastai provide to view the data in a DataLoaders?
Dls.one_batch() will give us the first batch of the training data
dfs.items will give us training and the testing items
where the dls is the dataloader which we've defined previously
10. What method does fastai provide to help you debug a DataBlock?
Datablock.summary() provides the summary how our dataset is getting generated
For cleaning the dataset we have interp.cleaner() which provides us functionality to go through the dataset and clean the dataset if necessary
11. Should you hold off on training a model until you have thoroughly cleaned your data?
No you should train the model first and then you can see where the model is getting confused most by determining that you can check the interp class and plot the most confuse values and also you can eliminate some pictures from the dataset
12. What are the two pieces that are combined into cross-entropy loss in PyTorch?
Nll_loss and log_sigmoid is combined in the cross_entropy loss
13. What are the two properties of activations that softmax ensures? Why is this important?
Softmax insures that the predicted probablity of the classes it applies standard exponential function and sum of the values to normalize it it amplifies results by large amount even if the data which we're passing to the result changes by the slight percentage
14. When might you want your activations to not have these two properties?
When there is binary classification we don't want the non linear activation functions we can just work with the sigmoid functions
15. Calculate the exp and softmax columns of <> yourself (i.e., in a spreadsheet, with a calculator, or in a notebook).
Exp function grows exponentially even for small change in the data
16. Why can't we use torch.where to create a loss function for datasets where our label can have more than two categories?
Since our data is not of the binary type we cant use the torch.where we have to use the cross entropy for finding out the losses
17. What is the value of $\log(-2)$? Why?
Not defined

18. What are two good rules of thumb for picking a learning rate from the learning rate finder?
Minimum loss learning rate / 10
and 2^{nd} when the last time learning rate started to decrease step!
19. What two steps does the fine_tune method do?
Freezing the layer
fit_one_cycle
unfreeze
fine_tune
20. In Jupyter Notebook, how do you get the source code for a method or function?
Function_name??
21. What are discriminative learning rates?
Learning rates are different for each layers the starting layers will have the smaller learning rate and ending layer will have larger learning rate we can pass lr_max = slice(start, end) for using discriminative learning rates
22. How is a Python slice object interpreted when passed as a learning rate to fastai?
The learning rate get applied to the first and last and the learning rate values for the in between layers is in multiples of that
23. Why is early stopping a poor choice when using 1cycle training?
If early stopping used then the training model will not have opportunity to test it one lower learning rate in turn keeps us away from getting the best result possible
24. What is the difference between resnet50 and resnet101?
Resnet 50 is having the 50 layers while the resnet 101 is having 101 layers
25. What does to_fp16 do?
It enables the mix precision in training which makes our models to train faster if the GPU used are latest (tesla t4 or tesla V100)