1. What is the concept of cyclical momentum?

Cyclical momentum refers to learning cycle or learning rate which is used for

1. faster training of the model
2. finding the optimal value of the parameter.

Cyclic learning rates have an impact on the model training process which is also known as “convergence”. Where we converge towards a point (global minima), by finding the optimal value of the parameters.

1. What callback keeps track of hyperparameter values (along with other data) during training?

We can keep track of hyper parameter values along with other data during training using AttrDict, text file like JSON, PyYAML or cfg files.

Eg :

Import yaml

from attrdict import AttrDict

with open(config\_path) as f:

config = yaml.load(f, Loader=yaml.BaseLoader)

cfg = AttrDict(config)

print(cfg.parameters.n\_cv\_splits)

3. In the color dim plot, what does one column of pixels represent?

In color dim plot, one column of pixels represents **RGB data value** (Red, Green, Blue).

1. In color dim, what does "poor teaching" look like? What is the reason for this?

1. Inaccurate or inconsistent use of color-related terminology, such as using hue, saturation, and value interchangeably or incorrectly.  
  
2. Failure to provide clear explanations of color concepts and principles, leaving students confused or uncertain about how to apply what they've learned.  
  
3. Lack of hands-on activities or visual aids to help students experience color interactions and mixing.  
  
4. Ignoring the cultural and psychological aspects of color that can affect how people perceive and respond to color.  
  
The reasons for poor teaching of color theory can vary, but they may include:  
  
1. Lack of training or expertise in color theory on the part of the teacher.  
  
2. Insufficient time or resources to devote to teaching color theory in depth.  
  
3. Reluctance to move beyond a narrow, preconceived notion of what color is and how it works.  
  
4. Limited opportunities for teacher professional development or peer collaboration to improve their teaching of color.

5. Does a batch normalization layer have any trainable parameters?

6. In batch normalization during preparation, what statistics are used to normalize? What about during the validation process?

7. Why do batch normalization layers help models generalize better?

When we normalize the image pixels it will range in between 0-1. By this way the dataset will be zero cantered. A good generalize model is created because weight updating will be faster and smother.

If we do analysis on variables measured in different scales, we might end up creating bias.

8.Explain between MAX POOLING and AVERAGE POOLING is number eight.

In Max Pooling we extract the maximum feature or the brighter pixels of an image. Whereas in Average Pooling we do average the pixel value and consider the same. Average pooling method smooths out the image, so the edges may not be identified.

9. What is the purpose of the POOLING LAYER?

We use pooling layer only to down sample the image pixel where default stride is 2. After pooling layer, we lose 50% of image pixels and 75% of the information. So, it is recommended to use pooling layer after 3-4 convolution layers.

10. Why do we end up with Completely CONNECTED LAYERS?

Fully connected layers also called as dense layers or hidden layers. When we use FC layer the number of parameters increase a lot. Which consumes a lot of time to train the model and in the closing leyer we do loose a lot of information. So we avoid FC layer in a network.

11. What do you mean by PARAMETERS?

In simple terms we can understand a parameter is a weight which is initialized to train the model and these parameters are trainable.

12. What formulas are used to measure these PARAMETERS?

The formula to measure the parameters:

no. of kernel\* ((kernel size) + bias)

let say we have applied 3x3 kernel 32 times where bias is 1

32,((3x3)+1) = 320