1. What is the concept of cyclical momentum?

**Ans : Learning rate might be the most important hyper parameter in deep learning, With One Cycle Policy and cyclic momentum ,**

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

**Ans : One traditional and popular way to perform hyperparameter tuning is by using an Exhaustive Grid Search from Scikit learn. This method tries every possible combination of each set of hyper-parameters. Using this method, we can find the best set of values in the parameter search space.**

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

**Ans : A pixel is the smallest block of color in a digital picture. The term is also used for the smallest block of color on your computer monitor.**

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

**Ans : But I am more interested whether simple neural network will be able to learn the pattern that will allow it to distinguish between colors**

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

**Ans : Consequently, batch normalization adds two trainable parameters to each layer, so the normalized output is multiplied by a “standard deviation” parameter (gamma) and add a “mean” parameter (beta).**

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

**Ans : Batch normalization is a technique to standardize the inputs to a network, applied to ether the activations of a prior layer or inputs directly. Batch normalization accelerates training, in some cases by halving the epochs or better, and provides some regularization, reducing generalization error.**

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

**Ans : Batch normalization is a technique to standardize the inputs to a network, applied to ether the activations of a prior layer or inputs directly. Batch normalization accelerates training, in some cases by halving the epochs or better, and provides some regularization, reducing generalization error.**

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

**Ans : Average pooling method smooths out the image and hence the sharp features may not be identified when this pooling method is used. Max pooling selects the brighter pixels from the image. It is useful when the background of the image is dark and we are interested in only the lighter pixels of the image.**

9. What is the purpose of the POOLING LAYER?

**Ans : Pooling layers are used to reduce the dimensions of the feature maps. Thus, it reduces the number of parameters to learn and the amount of computation performed in the network.**

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

**Ans : The aim of the Fully connected layer is to use the high-level feature of the input image produced by convolutional and pooling layers for classifying the input image into various classes based on the training dataset.**

11. What do you mean by PARAMETERS?

**Ans : A parameter is a quantity that influences the output or behavior of a mathematical object but is viewed as being held constant. Parameters are closely related to variables, and the difference is sometimes just a matter of perspective. For each value of the parameters, we get a different function.**

12. What formulas are used to measure these PARAMETERS?

**Ans : In the set of equations x = 2t + 1 and y = t2 + 2, t is called the parameter. As the parameter varies over a given domain of values, the set of solutions, or points (x, y), describes a curve in the plane.**