1. What is the concept of cyclical momentum?

The concept of cyclical momentum is a technique used in optimization algorithms, specifically in the context of training neural networks, to dynamically adjust the momentum parameter during the training process. It involves cyclically varying the momentum value over a predefined range during different training iterations.

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

The callback that keeps track of hyperparameter values, along with other data, during training is typically a custom callback that you define. You can create a custom callback class that inherits from the framework's base callback class and implement the desired functionality.

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

In the context of a color dim plot, one column of pixels represents the color values of a specific location or vertical strip of pixels in an image. The color dim plot visualizes the distribution and variation of color components across the image's width.

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

The color dim plot, "poor teaching" refers to a situation where the color dim plot does not effectively convey meaningful or informative patterns. It means that the plot fails to provide clear insights into the color distributions or variations within the image.

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

Yes, a batch normalization layer does have trainable parameters. The primary trainable parameters in a batch normalization layer are the scale and shift parameters.

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

During the training process, batch normalization normalizes the activations of a layer using the statistics calculated within each mini-batch. Specifically, the mean and variance of the activations within the mini-batch are computed and used to normalize the activations. The normalized activations are then scaled and shifted using the learned parameters (scale and shift) of the batch normalization layer.

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

Batch normalization layers help models generalize better by addressing the internal covariate shift problem and improving the stability and robustness of the training process.

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

Max Pooling:

In max pooling, a pooling window (often 2x2 or 3x3 in size) slides over the input feature map and extracts the maximum value within each window.

It down samples the input by selecting the maximum activation value, discarding the rest.

Average Pooling:

In average pooling, a pooling window slides over the input feature map and computes the average value within each window.

It down samples the input by taking the average activation value within each window.

1. What is the purpose of the POOLING LAYER?

The purpose of the pooling layer, also known as a pooling operation, in a convolutional neural network (CNN) is to perform down sampling or dimensionality reduction on the feature maps generated by the preceding convolutional layers. It helps to extract important features from the input while reducing the spatial dimensions of the feature maps.

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

We often end up with completely connected layers, also known as fully connected layers or dense layers, towards the end of a neural network architecture to enable high-level abstraction and facilitate classification or regression tasks.

1. What do you mean by PARAMETERS?

Parameters refer to the internal variables or weights that define the behavior and functionality of the model. Parameters are the values that the model learns during the training process to make predictions or perform a specific task.

1. What formulas are used to measure these PARAMETERS?
2. Loss Function
3. Gradient Calculation
4. Gradient Descent
5. Adaptive Learning Rate Algorithms
6. Stochastic Gradient Descent