1.What is the difference between TRAINABLE and NON-TRAINABLE PARAMETERS?

**Trainable Parameters vs. Non-trainable Parameters**:

* Trainable parameters are the weights and biases in a neural network that are adjusted during training to minimize the loss function. These parameters are learned from the data and are crucial for the model's ability to make predictions.
* Non-trainable parameters are fixed values or settings in the network that are not updated during training. They often include hyperparameters, fixed layers, or pre-trained embeddings. Non-trainable parameters are set before training and remain constant throughout.

2. In the CNN architecture, where does the DROPOUT LAYER go?

**Dropout Layer in CNN**:

* The dropout layer is typically added after one or more fully connected (dense) layers in a convolutional neural network (CNN). It helps prevent overfitting by randomly deactivating a fraction of neurons during training, effectively introducing a form of regularization. The exact placement of dropout layers can vary, but it's commonly added to fully connected layers or after flattening the output of convolutional layers.

3. What is the optimal number of hidden layers to stack?

**Optimal Number of Hidden Layers**:

* There is no one-size-fits-all answer to the optimal number of hidden layers in a neural network. It depends on the specific problem, data, and architecture. In practice, you can start with a simple architecture and gradually increase the complexity by adding hidden layers until you see diminishing returns in performance. Regularization techniques and larger datasets may allow deeper networks.

4. In each layer, how many secret units or filters should there be?

**Number of Neurons or Filters in Each Layer**:

* The number of neurons or filters in each layer depends on the complexity of the problem and the architecture. For convolutional layers, the number of filters can increase progressively as you move deeper into the network to capture more complex features. The number of neurons in fully connected layers can be determined based on the required model capacity and the number of features you want to learn.

5. What should your initial learning rate be?

**Initial Learning Rate**:

* The initial learning rate is a hyperparameter that you set before training. There's no one-size-fits-all value. Typically, a small learning rate (e.g., 0.1) is a reasonable starting point, and then you can use techniques like learning rate schedules or learning rate finders to fine-tune the learning rate.

6. What do you do with the activation function?

**Activation Function**:

* The activation function introduces non-linearity to the neural network. The choice of activation function depends on the specific task and architecture. Common choices include ReLU (Rectified Linear Unit), sigmoid, tanh, and others. The activation function transforms the weighted sum of inputs into an output that can capture complex relationships in the data.

7. What is NORMALIZATION OF DATA?

**Normalization of Data**:

* Normalization of data involves scaling the input features to a standardized range, typically with a mean of 0 and a standard deviation of 1. This process helps in training neural networks by ensuring that different features contribute equally to the learning process, preventing issues related to feature scales.

8. What is IMAGE AUGMENTATION and how does it work?

**Image Augmentation**:

* Image augmentation is a technique used to artificially increase the diversity of a dataset by applying various transformations to the images. This includes operations like rotation, flipping, cropping, and changes in brightness/contrast. Image augmentation is used to make the model more robust and to reduce overfitting by exposing it to different variations of the data during training.

9. What is DECLINE IN LEARNING RATE?

**Decline in Learning Rate**:

* The decline in learning rate, often referred to as learning rate annealing or learning rate scheduling, involves gradually reducing the learning rate during training. This can help the model converge more efficiently and achieve better performance.

10. What does EARLY STOPPING CRITERIA mean?

**Early Stopping Criteria**:

* Early stopping is a regularization technique that involves monitoring the model's performance on a validation set during training. Early stopping criteria determine when to stop training. It typically involves halting training when the validation performance stops improving or begins to degrade, preventing the model from overfitting the training data.