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

**Ans : Trainable parameters are the number of, well, trainable elements in your network; neurons that are affected by backpropagation. For example, for the Wx + b operation in each neuron, W and b are trainable – because they are changed by optimizers after backpropagation was applied for gradient computation. Nontrainable ones are e.g. Batch Normalization or when you lock/freeze layers e.g. during transfer learning.**

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

**Ans : Dropout can be used after convolutional layers (e.g. Conv2D) and after pooling layers (e.g. MaxPooling2D). Often, dropout is only used after the pooling layers, but this is just a rough heuristic. In this case, dropout is applied to each element or cell within the feature maps.**

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

**Ans : The number of hidden layers is highly dependent on the problem and the architecture of your neural network. You're essentially trying to Goldilocks your way into the perfect neural network architecture - not too big, not too small, just right. Generally, 1–5 hidden layers will serve you well for most problems.**

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

**Ans : The number of hidden neurons should be between the size of the input layer and the size of the output layer. The number of hidden neurons should be 2/3 the size of the input layer, plus the size of the output layer. The number of hidden neurons should be less than twice the size of the input layer.**

5. What should your initial learning rate be?

**Ans : A traditional default value for the learning rate is 0.1 or 0.01, and this may represent a good starting point on your problem.**

6. What do you do with the activation function?

**Ans : an activation function is a function that is added into an artificial neural network in order to help the network learn complex patterns in the data. When comparing with a neuron-based model that is in our brains, the activation function is at the end deciding what is to be fired to the next neuron.**

7. What is NORMALIZATION OF DATA?

**Ans : Data normalization is generally considered the development of clean data. Diving deeper, however, the meaning or goal of data normalization is twofold:**

* **Data normalization is the organization of data to appear similar across all records and fields.**
* **It increases the cohesion of entry types leading to cleansing, lead generation, segmentation, and higher quality data.**

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

**Ans : Image augmentation is a technique of altering the existing data to create some more data for the model training process. In other words, it is the process of artificially expanding the available dataset for training a deep learning model.**

9. What is DECLINE IN LEARNING RATE?

**Ans : Perhaps the simplest learning rate schedule is to decrease the learning rate linearly from a large initial value to a small value. This allows large weight changes in the beginning of the learning process and small changes or fine-tuning towards the end of the learning process.**

10. What does EARLY STOPPING CRITERIA mean?

**Ans : In machine learning, early stopping is a form of regularization used to avoid overfitting when training a learner with an iterative method, such as gradient descent. Early stopping rules provide guidance as to how many iterations can be run before the learner begins to over-fit.**