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

* Trainable parameters are parameters of a machine learning model that can be optimized through training, while non-trainable parameters are those that cannot be changed during training. Trainable parameters typically include the weights and biases of a model, while non-trainable parameters might include the structure of the model or other hyperparameters.

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

* In a convolutional neural network (CNN), the dropout layer is typically placed after the fully connected layers. This is because the fully connected layers have a lot of parameters, and using a dropout layer after these layers can help prevent overfitting. The dropout layer randomly drops a specified fraction of the input units during training, which can help prevent overfitting and improve the generalizability of the model.

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

* In general, as the complexity of the problem increases, the number of hidden layers in the network should also increase in order to capture the necessary patterns in the data. However, adding too many hidden layers can make the model too complex and can lead to overfitting. It is generally recommended to start with a simple model and then increase the number of hidden layers as needed. Experimenting with different numbers of hidden layers and using techniques such as cross-validation can help you find the best number of hidden layers for your particular problem.

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

* The number of units or filters in each layer of a neural network is a hyperparameter that can affect the performance of the model. There is no single optimal number of units or filters that will work for all problems, and the appropriate number to use can depend on a variety of factors such as the complexity of the problem and the amount of data available for training. In general, as the complexity of the problem increases, the number of units or filters in each layer should also increase in order to capture the necessary patterns in the data. However, adding too many units or filters can make the model too complex and can lead to overfitting. It is generally recommended to start with a simple model and then increase the number of units or filters as needed. Experimenting with different numbers of units or filters and using techniques such as cross-validation can help you find the best number for your particular problem.

5. What should your initial learning rate be?

* The initial learning rate is a hyperparameter that determines the step size at which the model is updated during training. The appropriate initial learning rate can vary depending on the specific problem you are trying to solve and the characteristics of your dataset. In general, a larger learning rate can allow the model to learn faster, but it can also make the model more unstable and cause it to converge to a suboptimal solution. On the other hand, a smaller learning rate can make the model more stable, but it can also cause the training process to be slow and require more epochs to achieve good performance. It is generally recommended to start with a moderate learning rate and then adjust it as needed based on the performance of the model on the validation set. Experimenting with different learning rates and using techniques such as learning rate schedules can help you find the best initial learning rate for your particular problem.

6. What do you do with the activation function?

* The activation function is typically a non-linear function such as the sigmoid, tanh, or ReLU function. The choice of activation function can have a significant impact on the performance of the model, and different activation functions can work better for different types of problems.

7. What is NORMALIZATION OF DATA?

* Normalization of data is a method used to scale the values of the data to a specific range. This is typically done to bring all the values of the data into the same range, which can make it easier for the model to learn and can improve the performance of the model. There are several different methods for normalizing data, such as min-max scaling, z-score normalization, and decimal scaling. The appropriate method to use can depend on the characteristics of the data and the specific problem you are trying to solve. Normalizing the data can also help to reduce the impact of outliers in the data and can make the model more robust. It is generally recommended to normalize the data before training a machine learning model in order to improve its performance.

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

* Image augmentation is a technique used to artificially increase the size of a training dataset by creating modified versions of the existing data. This is done by applying various transformations to the images, such as rotating, shifting, scaling, or flipping the images. These transformed images can be used in addition to the original images to train a machine learning model. Image augmentation can help to reduce overfitting, improve the generalizability of the model, and can make the model more robust. It is commonly used in computer vision applications where the amount of available training data may be limited. To perform image augmentation, a library or framework that provides this functionality can be used, such as TensorFlow, Keras, or PyTorch. These libraries typically provide a set of predefined image augmentation techniques that can be easily applied to the data.

9. What is DECLINE IN LEARNING RATE?

* Decline in learning rate refers to a strategy used to reduce the step size at which the model is updated during training. This can be done by decreasing the learning rate over time, either at a fixed rate or according to a schedule. Declining the learning rate during training can help the model to converge to a better solution and can improve the performance of the model. It can also help to prevent the model from getting stuck in a suboptimal solution or from overfitting to the training data. There are several different methods for declining the learning rate, such as using a fixed learning rate schedule or using a learning rate decay function. The appropriate method to use can depend on the specific problem you are trying to solve and the characteristics of the dataset. It is generally recommended to experiment with different learning rate schedules and find the one that works best for your particular problem.

What does EARLY STOPPING CRITERIA mean?

* Early stopping criteria is a technique used in machine learning to determine the appropriate time to stop training a model. It involves monitoring the performance of the model on a separate validation set at regular intervals during training, and stopping the training process once the performance on the validation set stops improving or starts to deteriorate. This can prevent the model from overfitting, which is when a model performs well on the training data but poorly on new, unseen data. Early stopping criteria can help to improve the generalization of the model and make it more effective on a variety of tasks.