**Assignment - 06**

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

Ans: Difference between TRAINABLE and NON-TRAINABLE PARAMETERS:

Trainable parameters are the weights and biases in a neural network that are updated during the training process to minimize the loss function.

Non-trainable parameters are constants or fixed values that are not updated during training. They include parameters in layers such as batch normalization, dropout layers, and activation functions, as well as pretrained weights in transfer learning.

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

Ans: Placement of DROPOUT LAYER in CNN architecture:

The dropout layer is typically placed after the convolutional and pooling layers, and before the fully connected layers in a CNN architecture.

It randomly drops a fraction of neurons during training, helping prevent overfitting by reducing co-adaptation between neurons.

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

Ans: Optimal number of hidden layers to stack

The optimal number of hidden layers in a neural network depends on various factors such as the complexity of the problem, the size of the dataset, and computational resources.

There is no one-size-fits-all answer, but common architectures often have 2-5 hidden layers.

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

Ans: Number of units or filters in each layer

The number of units or filters in each layer depends on the complexity of the problem, the size of the dataset, and the architectural choices.

It is often determined through experimentation and tuning to achieve the best performance.

5. What should your initial learning rate be?

Ans: The initial learning rate is a hyperparameter that determines the size of the steps taken during optimization.

Common initial learning rates range from 0.001 to 0.1, but the optimal value depends on the problem, architecture, and optimization algorithm.

6. What do you do with the activation function?

Ans: Activation functions introduce non-linearity to the neural network, enabling it to learn complex patterns and relationships in the data.

The choice of activation function depends on the problem, but common options include ReLU, sigmoid, tanh, and softmax.

7. What is NORMALIZATION OF DATA?

Ans: Normalization is the process of scaling input features to a similar range to ensure that they contribute equally to the model's learning process.

Common normalization techniques include z-score normalization (standardization) and min-max scaling.

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

Ans: Image augmentation is a technique used to artificially increase the size of the training dataset by applying transformations such as rotation, flipping, scaling, and cropping to input images.

It helps improve the generalization and robustness of the model by exposing it to a wider variety of training examples.

9. What is DECLINE IN LEARNING RATE?

Ans: Decline in learning rate refers to the reduction of the learning rate during training, often implemented using learning rate schedules or techniques like learning rate decay.

It helps stabilize the training process, prevent overshooting, and fine-tune the model parameters towards convergence.

10 What does EARLY STOPPING CRITERIA mean?

Ans: Early stopping criteria is a regularization technique used to prevent overfitting by stopping the training process when the performance on a validation set stops improving.

It involves monitoring a performance metric (e.g., validation loss or accuracy) and stopping training if the metric fails to improve for a specified number of epochs (patience).