1. What is the COVARIATE SHIFT Issue, and how does it affect you?

2. What is the process of BATCH NORMALIZATION?

3. Using our own terms and diagrams, explain LENET ARCHITECTURE.

4. Using our own terms and diagrams, explain ALEXNET ARCHITECTURE.

5. Describe the vanishing gradient problem.

6. What is NORMALIZATION OF LOCAL RESPONSE?

7. In AlexNet, what WEIGHT REGULARIZATION was used?

8. Using our own terms and diagrams, explain VGGNET ARCHITECTURE.

9. Describe VGGNET CONFIGURATIONS.

10. What regularization methods are used in VGGNET to prevent overfitting?

Answer:

1. COVARIATE SHIFT refers to the change in the distribution of input features or covariates during training and testing, which can cause a mismatch between the training and testing datasets. This can lead to poor generalization performance of the model and reduced accuracy on the testing set.
2. BATCH NORMALIZATION is a technique used to normalize the activations of the previous layer for each training mini-batch. This helps to reduce the internal covariate shift and stabilize the learning process. It involves computing the mean and variance of the activations over the mini-batch, and normalizing the activations to have zero mean and unit variance. It then applies a scaling and shifting parameter to the normalized activations to allow the model to learn the optimal scale and shift for each activation.
3. LENET ARCHITECTURE is a type of convolutional neural network (CNN) that was developed in the early 1990s by Yann LeCun for the purpose of recognizing handwritten digits. The architecture consists of two sets of convolutional and pooling layers followed by a fully connected layer. The first set of convolutional and pooling layers is used to extract local features from the input image, while the second set is used to extract global features. The fully connected layer is then used for classification.
4. ALEXNET ARCHITECTURE is a deep convolutional neural network (CNN) architecture that was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton for the ImageNet Large Scale Visual Recognition Challenge in 2012. The architecture consists of five convolutional layers, some of which are followed by max pooling layers, and three fully connected layers. It uses the rectified linear unit (ReLU) activation function and dropout regularization to prevent overfitting.
5. The VANISHING GRADIENT PROBLEM occurs in deep neural networks when the gradient of the loss function becomes very small or zero during backpropagation, making it difficult for the model to learn. This is due to the multiplication of small gradients in the chain rule of derivatives, which causes the gradient to "vanish" as it propagates backwards through the network. As a result, the earlier layers of the network may not receive enough information to learn meaningful features.
6. NORMALIZATION OF LOCAL RESPONSE (NLR) is a technique used in convolutional neural networks (CNNs) to enhance the contrast of local features. It involves normalizing the response of each neuron in a local neighborhood of the input image by dividing it by the sum of the squared responses of all neurons in that neighborhood. This amplifies the response of neurons with high activity relative to their neighbors and improves the discriminability of local features.
7. In ALEXNET, WEIGHT REGULARIZATION is used to prevent overfitting by adding a penalty term to the loss function that penalizes large weights. Specifically, the L2 regularization term is added to the loss function, which is proportional to the sum of the squares of all the weights in the network.
8. VGGNET ARCHITECTURE is a type of convolutional neural network (CNN) that was developed by the Visual Geometry Group (VGG) at the University of Oxford. It consists of 16 or 19 layers, which are organized into blocks of convolutional layers followed by max pooling layers. The convolutional layers use small 3x3 filters, which helps to reduce the number of parameters in the network. The architecture also uses dropout regularization and batch normalization to prevent overfitting.
9. VGGNet has three configurations: VGG-11, VGG-16, and VGG-19. These configurations differ in the number of convolutional layers and the size of the fully connected layers. VGG-11 has 11 layers (8 convolutional and 3 fully connected), VGG-16 has 16 layers (13 convolutional and 3 fully connected), and VGG-19 has 19 layers (16 convolutional and 3 fully connected).
10. VGGNet uses dropout regularization to prevent overfitting. Dropout randomly sets some activations to zero during training, which helps to prevent the network from relying too much on any one activation. Additionally, VGGNet uses weight decay regularization, which adds a penalty term to the loss function based on the magnitude of the weights in the network. This encourages the network to use smaller weights, which can help to prevent overfitting.