**Assignment\_7**

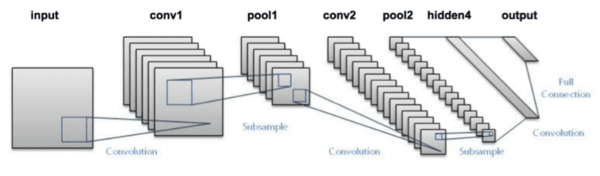
1.What is the COVARIATE SHIFT Issue, and how does it affect you?

**Ans: Covariate shift refers to the change in the distribution of the input variables present in the training and the test data. It is the most common type of shift and it is now gaining more attention as nearly every real-world dataset suffers from this problem.**

2. What is the process of BATCH NORMALIZATION?

**Ans: Batch normalization, is a process to make neural networks faster and more stable through adding extra layers in a deep neural network. The new layer performs the standardizing and normalizing operations on the input of a layer coming from a previous layer.**

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

Ans: 

**The network has 5 layers with learnable parameters and hence named Lenet-5. It has three sets of convolution layers with a combination of average pooling. After the convolution and average pooling layers, we have two fully connected layers. At last, a Softmax classifier which classifies the images into respective class.**

**The input to this model is a 32 X 32 grayscale image hence the number of channels is one.**

**We then apply the first convolution operation with the filter size 5X5 and we have 6 such filters. As a result, we get a feature map of size 28X28X6. Here the number of channels is equal to the number of filters applied.**

**After the first pooling operation, we apply the average pooling and the size of the feature map is reduced by half. Note that, the number of channels is intact.**

**Next, we have a convolution layer with sixteen filters of size 5X5. Again the feature map changed it is 10X10X16. The output size is calculated in a similar manner. After this, we again applied an average pooling or subsampling layer, which again reduce the size of the feature map by half i.e 5X5X16.**

**Then we have a final convolution layer of size 5X5 with 120 filters. As shown in the above image. Leaving the feature map size 1X1X120. After which flatten result is 120 values.**

**After these convolution layers, we have a fully connected layer with eighty-four neurons. At last, we have an output layer with ten neurons since the data have ten classes.**

**Here is the final architecture of the Lenet-5 model.**

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

**Ans: AlexNet. The architecture consists of eight layers: five convolutional layers and three fully-connected layers. But this isn’t what makes AlexNet special; these are some of the features used that are new approaches to convolutional neural networks:**

* **ReLU Nonlinearity. AlexNet uses Rectified Linear Units (ReLU) instead of the tanh function, which was standard at the time. ReLU’s advantage is in training time; a CNN using ReLU was able to reach a 25% error on the CIFAR-10 dataset six times faster than a CNN using tanh.**
* **Multiple GPUs. Back in the day, GPUs were still rolling around with 3 gigabytes of memory (nowadays those kinds of memory would be rookie numbers). This was especially bad because the training set had 1.2 million images. AlexNet allows for multi-GPU training by putting half of the model’s neurons on one GPU and the other half on another GPU. Not only does this mean that a bigger model can be trained, but it also cuts down on the training time.**
* **Overlapping Pooling. CNNs traditionally “pool” outputs of neighboring groups of neurons with no overlapping. However, when the authors introduced overlap, they saw a reduction in error by about 0.5% and found that models with overlapping pooling generally find it harder to overfit.**

**The Overfitting Problem. AlexNet had 60 million parameters, a major issue in terms of overfitting. Two methods were employed to reduce overfitting:**

* **Data Augmentation. The authors used label-preserving transformation to make their data more varied. Specifically, they generated image translations and horizontal reflections, which increased the training set by a factor of 2048. They also performed Principle Component Analysis (PCA) on the RGB pixel values to change the intensities of RGB channels, which reduced the top-1 error rate by more than 1%.**
* **Dropout. This technique consists of “turning off” neurons with a predetermined probability (e.g. 50%). This means that every iteration uses a different sample of the model’s parameters, which forces each neuron to have more robust features that can be used with other random neurons. However, dropout also increases the training time needed for the model’s convergence.**

**The Results. On the 2010 version of the ImageNet competition, the best model achieved 47.1% top-1 error and 28.2% top-5 error. AlexNet vastly outpaced this with a 37.5% top-1 error and a 17.0% top-5 error. AlexNet is able to recognize off-center objects and most of its top five classes for each image are reasonable. AlexNet won the 2012 ImageNet competition with a top-5 error rate of 15.3%, compared to the second place top-5 error rate of 26.2%.**

5. Describe the vanishing gradient problem.

**Ans: It describes the situation where a deep multilayer feed-forward network or a recurrent neural network is unable to propagate useful gradient information from the output end of the model back to the layers near the input end of the model.**

6. What is NORMALIZATION OF LOCAL RESPONSE?

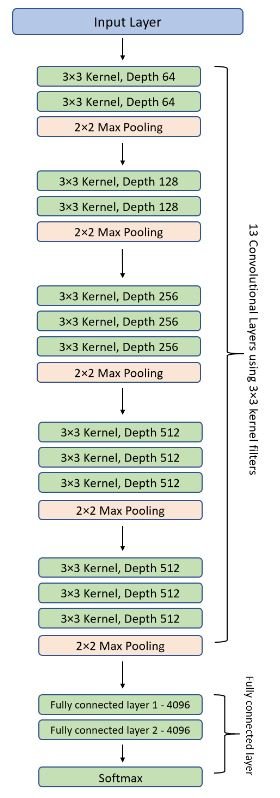
**Ans: Local Response Normalization (LRN) was first introduced in AlexNet architecture where the activation function used was ReLU as opposed to the more common tanh and sigmoid at that time. Apart from the reason mentioned above, the reason for using LRN was to encourage lateral inhibition.**

7. In AlexNet, what WEIGHT REGULARIZATION was used?

**Ans: Weight regularization provides an approach to reduce the overfitting of a deep learning neural network model on the training data and improve the performance of the model on new data, such as the holdout test set.**

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

**Ans: VGG stands for Visual Geometry Group; it is a standard deep Convolutional Neural Network (CNN) architecture with multiple layers. The “deep” refers to the number of layers with VGG-16 or VGG-19 consisting of 16 and 19 convolutional layers. The VGG architecture is the basis of ground-breaking object recognition models.**



9. Describe VGGNET CONFIGURATIONS.

**Ans: VGG incorporates 1x1 convolutional layers to make the decision function more non-linear without changing the receptive fields. The small-size convolution filters allows VGG to have a large number of weight layers; of course, more layers leads to improved performance.**

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

**Ans: Regularization. Regularization optimizes a model by penalizing complex models, therefore minimizing loss and complexity. Thus this forces our neural network to be simpler. Here we will use an L2 regularizer, as it is the most common and is more stable than an L1 regularizer**.