1. Explain about Nesterov accelerated gradient descent and RMSProp optimizers in details.
2. Explain about dimensionality reduction using Autoencoders with an example.
3. Explain the concept of skip-gram model with an example.
4. You come up with a CNN classifier for 3 classes. For each layer, calculate the output size and number of parameters associated with each layer.

Layer	Activation map size	No of parameters
input	128*128*3	0
Conv -5-16		
MaxPool-2		
Conv-3-32		
MaxPool-2		
Conv-3-64		
MaxPool-2		
FC-1(128)		
FC-2(25)		

The notation follows the convention:
• CONV-K-N denotes a convolutional layer with N filters, each of them of size K x K. Padding and stride parameters are always 0 and 1 respectively.
• POOL-K indicates a K x K pooling layer with stride K and padding 0.
FC(X) indicates X number of neurons.

5. Write in detail about principal component analysis with an example
6. What are the limitations of the vennila gradient descent algorithm?
(i) Explain the techniques to overcome overfitting in deep neural network models.
(ii) Compare PCA and Auto encoders
6. Consider a CNN with the following layers stacked 1-convolution, 1-activation(ReLu), 1-Avg pooling, 1-flattening and 1-dense layer with 3 neurons and softmax function. Show the processing and output of each layer for the following input data.

3	3	3	1	0
0	0	1	3	1
3	1	2	2	3
2	0	0	2	2
2	0	0	0	1

Kernel filter 3*3

0	1	2
-1	2	0
0	-1	2

7. Write in detail about AlexNet and Google Net with their architectures and highlight the difference between them.
8. Explain how the following concepts help to overcome overfitting. (a)L2 regularization (b)Data augmentation
9. What is dimensionality deduction? Illustrate the concept of SVD.
10. Calculate CNN operation using below
11. Using 5*5 input data calculate using kernel filter

3	3	3	1	0
0	0	1	3	1
3	1	2	2	3
2	0	0	2	2
2	0	0	0	1

Kernel filter 3*3

0	1	2
2	2	0
0	1	2

12. Use input to calculate below operation in CNN
1)stride
2)padding
3)max pooling
4)avg pooling

12. (i)What are the limitations of vanilla gradient descent learning algorithm?
(ii)How can you overcome the same by the following optimization techniques? Detail the theory and formulas for each optimization technique in detail?
(1)Adagrad
(2)Nestrov
(3)Adam

13 Explain four concepts that help to overcome **overfitting** in deep neural networks? Explain each of them in detail with neat diagram
14 Explain the intuition behind Resnet and detail the architecture with a neat diagram

15. You come up with a CNN classifier for 3 classes. For each layer, calculate the number of weights, number of biases and the size of the associated feature maps.

Explain in detail with proper steps and calculations:

Layer	Activation map	No of weights	No of biases
input	256*256*3	0	0
Conv -9-32			
Pool-2			
Conv-5-64			
Pool-2			
Conv-5-64			
Pool-2			
FC-1			
FC-2			

The notation follows the convention:

• CONV-K-N denotes a convolutional layer with N filters, each of them of size K x K.Padding and stride parameters are always 0 and 1 respectively.
• POOL-K indicates a K x K pooling layer with stride K and padding 0.

Given an input RGB image of size 32X32, a convolution with 8 3x3 filters are applied. Stride and padding of convolution is 1 and 2 respectively. What is the size of output feature map?
Write the algorithm for a “Mini-batch gradient descent algorithm”

Question Bank From 4th and 5th UNIT

UNIT 4

RNN with an example

What is the use of GRU? Compare with LSTM.

Explain in detail VGG16 architecture for image classification.

LSTM architecture and build an LSTM network for named entity recognition.

Transfer Learning, need of transfer learning, advantages and disadvantages of transfer learning

exploding gradients and Vanishing Gradients and Illustrate how LSTM helps to solve the Gradients Problems.

DenseNet architecture. What problems are solved by DenseNet ?

UNIT 5

Machine Translation

Generative adversial network (GAN).

Restricted Boltzmann Machines (RBM).

Attention mechanism for prediction purposes

RBM training with Block Gibbs sampling

hierarchical attention in deep learning

the encoder and decoder Models and list out the applications of Encoder and decoder Models.

language modeling with its types and real time examples.

Explain Generative adversarial network (GAN) with a neat diagram and list out the types of GAN.