Deep Learning assignment

Section A: Multiple-Choice Questions

1. What is the primary function of an activation function in a neural network?
A) To initialize the weights of the network
B) To introduce non-linearity into the model
C) To reduce the size of the dataset
D) To optimize the learning rate

2. Which of the following is a commonly used loss function for classification tasks?
A) Mean Squared Error (MSE)
B) Cross-Entropy Loss
C) Hinge Loss
D) Huber Loss

3. What does the term 'overfitting' refer to in deep learning?
A) When the model performs well on training data but poorly on unseen data
B) When the model performs poorly on both training and unseen data
C) When the model has too few parameters to learn the data
D) When the model uses too much computational power

4. Which optimization algorithm is commonly used in training deep learning models?
A) Gradient Descent
B) Newton's Method
C) Simulated Annealing
D) Genetic Algorithm

5. What is the purpose of dropout in a neural network?
A) To increase the size of the dataset
B) To prevent overfitting by randomly deactivating neurons during training
C) To speed up the training process
D) To initialize the weights of the network

6. What is the role of a convolutional layer in a Convolutional Neural Network (CNN)?
A) To reduce the dimensionality of the input data
B) To extract spatial features from the input data
C) To optimize the learning rate
D) To initialize the weights of the network

7. Which of the following techniques helps prevent overfitting in deep learning?
A) Batch normalization
B) Dropout
C) Data augmentation
D) All of the above

8. What is the vanishing gradient problem in deep learning?
A) When gradients become too small during backpropagation, hindering learning

B) When gradients become too large during backpropagation, causing instability
C) When the model fails to converge due to poor initialization
D) When the model overfits the training data

9. What is the purpose of batch normalization in deep learning?
A) To reduce overfitting by randomly deactivating neurons
B) To normalize the input data to each layer, improving training stability
C) To increase the size of the dataset
D) To optimize the learning rate

10. Which architecture is commonly used for image classification tasks?
A) Recurrent Neural Network (RNN)
B) Convolutional Neural Network (CNN)
C) Generative Adversarial Network (GAN)
D) Transformer

(20 points)

Section B: Critical Thinking Questions

11. How would you modify a neural network to improve generalization without significantly increasing training time?

12. If a deep learning model is underfitting despite having sufficient data, what possible factors could be causing this, and how would you address them?

13. You are training a deep learning model, but you notice that with a high learning rate, the loss fluctuates wildly, while with a very low learning rate, the model converges too slowly. How would you design a learning rate schedule to optimize training efficiency while ensuring stable convergence?

14. Transfer learning can speed up model training, but it may introduce unwanted biases from the original dataset. How would you ensure a transferred model is well-adapted to a new domain while avoiding inherited bias?

15. Deep learning models often require large amounts of labeled data for training. How would you approach improving model performance when labeled data is scarce or difficult to obtain?

(50 points)

Section C: CNN Activation Shape & Learnable Parameters Calculation

16. Consider a CNN with an input image of shape (32, 32, 3). The first convolutional layer has 16 filters of size (3,3), stride 1, and padding is 'valid'. What is the shape of the output feature map?

17. If a CNN has a max pooling layer with pool size (2,2) and stride 2 applied after the first convolutional layer, what will be the shape of the feature map after pooling?

18. A CNN consists of the following layers:

Neural Networks

Input: (256, 256, 3)

Conv Layer: 32 filters (3,3), stride 1, padding 'valid'

Max Pooling: pool size (2,2), stride 2

Conv Layer: 64 filters (3,3), stride 1, padding 'same'

Flatten Layer

Calculate the total number of learnable parameters in the CNN.

(30 points)

Good luck !!!