School of Engineering and Applied Science (SEAS), Ahmedabad University

Name: _____ Roll no.: _____

B. Tech. (Computer Science and Engineering)
CSE602 Deep Learning: Mid Term Exam (Winter Semester: 2025-26); Marks: 30
February 24, 2025, Tuesday, 05:00 PM

Instructions:

1 All questions are compulsory.
2 Each question contributes one mark.
3 All acronyms carry their usual meaning.
4 Assume suitable data wherever necessary and specify it clearly in your answers.

Answer the Questions

1. What may increase model complexity in deep networks?
   A. Pooling layers
   B. Activation function
   C. Larger filters
   D. Batch normalisation

2. CNNs are best suited for what data types?
   A. Text
   B. Audio
   C. Image
   D. Tabular data

3. What does flattening do in deep networks?
   A. Compresses model parameters
   B. Removes noise
   C. Reduces overfitting
   D. Converts n-D data into 1-D

4. What does stride in convolution operation controls?
   A. Depth
   B. Output accuracy
   C. Movement of the filter
   D. Loss value

5. Which optimizer can be used in training a deep neural network?
   A. RMSProp

B. Adam
C. Stochastic gradient descent
D. All

6. What does Batch size refers to in deep neural networks?
A. Size of the input image
B. Number of epochs
C. Number of samples processed before updating weights
D. Number of layers

7. Suppose a dataset has 100 data samples and the deep neural network has passed the full dataset 5 times. What will be the value of the epoch parameter?
A. 100
B. 500
C. 5
D. 20

8. Suppose a deep network is classifying images of true or fake apples. Which loss function is suitable for it?
A. Categorical cross-entropy
B. Hinge loss
C. Binary cross-entropy
D. Mean squared error

9. What does early stopping method do?
A. Trains a model faster
B. Avoids underfitting
C. Prevents overfitting
D. Save memory

10. Which of these is a regularization technique?
A. Backpropagation
B. L2
C. Softmax
D. One-hot encoding

11. In a deep neural network training, what does learning rate parameter controls?
A. Loss function
B. How fast model updates weights
C. Input data size

D. Training duration

12. What does hyperparameter tuning involve?
    A. Feature scaling
    B. Adjusting model parameters manually
    C. Data augmentation
    D. Model evaluation

13. Which tool is used to visualize training metrics?
    A. PyTorch
    B. TensorBoard
    C. Pandas
    D. NumPy

14. In GANs, what does the generator aim to?
    A. Classify real data
    B. Discriminate data
    C. Fool the discriminator
    D. Train CNNs

15. What is the purpose of dropout regularization in deep learning?
    A. To reduce overfitting
    B. To increase the model's capacity
    C. To improve the training speed
    D. To handle imbalanced datasets

16. What is the purpose of a loss function in deep learning?
    A. To measure the difference between predicted and actual values
    B. To guide the training process
    C. To compute the gradients for updating the network's parameters
    D. All of the above

17. Which deep learning technique is used for unsupervised feature learning?
    A. Autoencoders
    B. Convolutional Neural Networks
    C. Recurrent Neural Networks
    D. Reinforcement Learning

18. Which deep learning technique is used for learning latent representations from unlabelled data?

A. Self-supervised learning
B. Reinforcement Learning
C. Transfer Learning
D. Unsupervised Learning

19. Suppose you have a 3x3 image with all 1s. If the kernel is [[1 0] [0 1]], what could be the convolved feature? Assume suitable padding and stride values.
A. [[1 1][1 1]]
B. [[2 1] [1 2]]
C. [[2 2] [2 2]]
D. [[0 1] [0 1]]

20. Suppose you have a 4x4 image with all 3s. If you perform max pooling operation on it with stride 2, what will be the total number of elements in the outcome?
A. 2
B. 3
C. 4
D. 6

21. Assume a grey scale image I of size 16x16. What will be the value range in I after performing I/255.0?
A. -1 to 1
B. 0 to 255
C. 0 to 1
D. -128 to 127

22. Weights learnerd by hidden layers cannot be visualised, but the final outcome can be visualised. Select True/False.
A. True
B. False

23. What type of modelling approach can be adapted for sentiment analysis?
A. One-to-many RNN
B. Many-to-many RNN
C. Many-to-one RNN
D. One-to-one RNN

24. What type of sequence modelling approach can be adapted for language translation?
A. One-to-many RNN
B. Many-to-many RNN

C. Many-to-one RNN
D. One-to-one RNN

25. What among the below is not a gate in the long short-term memory network?
A. Input
B. Output
C. Memory
D. Forget

26. LSTM is a light-weight version of GRU. Select True/False.
A. True
B. False

27. What optimiser is suitable for the fast and rapid convergence?
A. Gradient descent
B. Stochastic gradient descent
C. AdaGrad
D. Adam

28. What is not the technique to prevent vanishing gradients?
A. Batch normalisation
B. ReLU activation for hidden layers
C. Skip connections
D. None

29. What is not the sign for exploding gradients?
A. Fluctuating loss function
B. NaN during the training process
C. Weights converging to zero during the training
D. None

30. What may cause overfitting?
A. Model complexity
B. High variance and low bias in the data
C. K-fold cross-validation
D. Dropout

All the best
