



**SNS COLLEGE OF ENGINEERING**  
**Kurumbapalayam (Po), Coimbatore – 641 107**  
**AN AUTONOMOUS INSTITUTION**



**Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai**  
**DEPARTMENT OF ARTIFICIAL INTELLIGENCE AND DATA SCIENCE**  
**PUZZLE QUESTION**  
**DEEP LEARNING**

## 1. Linear (PCA, LDA) and Manifolds

- **Puzzle:** Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) both reduce the dimensionality of data. However, which of the following is a key difference between PCA and LDA?
- **Options:**
  - A) PCA maximizes variance, whereas LDA maximizes class separability.
  - B) PCA is supervised, whereas LDA is unsupervised.
  - C) PCA is used for classification, whereas LDA is used for regression.
  - D) PCA focuses on distance metrics, whereas LDA does not.
- **Answer:** A) PCA maximizes variance, whereas LDA maximizes class separability.

## 2. Manifolds

- **Puzzle:** In the context of machine learning, data often lies on a "manifold" in a high-dimensional space. What is the main benefit of identifying a manifold structure in data?
- **Options:**
  - A) It helps reduce computational complexity by reducing the number of features needed for training.
  - B) It leads to faster convergence of deep learning models.
  - C) It allows us to perform better in anomaly detection tasks.
  - D) It increases the size of the dataset for better training.
- **Answer:** A) It helps reduce computational complexity by reducing the number of features needed for training.

## 3. Metric Learning

- **Puzzle:** In metric learning, the goal is to learn a distance function that can compare pairs of examples. Which of the following is a typical application of metric learning?
- **Options:**
  - A) Image captioning
  - B) Sentiment analysis
  - C) Face verification and recognition
  - D) Machine translation
- **Answer:** C) Face verification and recognition

## 4. Autoencoders and Dimensionality Reduction in Networks

- **Puzzle:** Autoencoders are used for dimensionality reduction. In an autoencoder network, the "bottleneck" layer is designed to:
- **Options:**
  - A) Maximize the reconstruction error
  - B) Force the model to compress the input data into a lower-dimensional representation
  - C) Increase the number of neurons to enhance the representation power
  - D) Expand the data before reconstruction
- **Answer:** B) Force the model to compress the input data into a lower-dimensional representation

## 5. Introduction to ConvNet

- **Puzzle:** Convolutional Neural Networks (ConvNets) are widely used in image recognition. What is the primary function of the convolutional layer in a ConvNet?
- **Options:**
  - A) To apply non-linear activations to the output of the previous layer
  - B) To perform matrix multiplication with input data
  - C) To detect low-level features such as edges or textures
  - D) To reduce the dimensionality of the data
- **Answer:** C) To detect low-level features such as edges or textures

## 6. Architectures – AlexNet, VGG, Inception, ResNet

- **Puzzle:** Which of the following network architectures is known for introducing the concept of "skip connections," which help alleviate the vanishing gradient problem in deep networks?
- **Options:**
  - A) AlexNet
  - B) VGG
  - C) Inception
  - D) ResNet
- **Answer:** D) ResNet

## 7. AlexNet

- **Puzzle:** AlexNet, a deep convolutional neural network, was a significant breakthrough in computer vision. Which of the following was a major reason for its success in the 2012 ImageNet competition?
- **Options:**

- A) Its ability to work with very small datasets
- B) The use of large filter sizes in the convolution layers
- C) The use of ReLU activation functions, which helped with training speed
- D) Its use of dropout regularization
- **Answer:** C) The use of ReLU activation functions, which helped with training speed

## 8. VGG

- **Puzzle:** VGG is another popular ConvNet architecture. What is the key characteristic of the VGG network that distinguishes it from other networks?
- **Options:**
  - A) The use of very large fully connected layers
  - B) The use of small 3x3 convolutional filters stacked in deep layers
  - C) The use of 1x1 convolutions to reduce the feature maps
  - D) The inclusion of residual connections
- **Answer:** B) The use of small 3x3 convolutional filters stacked in deep layers

## 9. Training a ConvNet: Weights Initialization

- **Puzzle:** When training a ConvNet, weight initialization is crucial. Which of the following methods helps prevent the vanishing and exploding gradient problem in deep networks?
- **Options:**
  - A) Randomly initializing weights to large values
  - B) Initializing all weights to zero
  - C) Using methods like Xavier or He initialization
  - D) Initializing weights based on the mean of the input data
- **Answer:** C) Using methods like Xavier or He initialization

## 10. Batch Normalization and Hyperparameter Optimization

- **Puzzle:** Batch normalization has become a standard practice in training deep networks. What is the primary effect of batch normalization on the training process?
- **Options:**
  - A) It increases the learning rate.
  - B) It normalizes the inputs of each layer, reducing the impact of poor weight initialization.
  - C) It forces the network to converge faster by reducing the number of training epochs.
  - D) It increases the number of parameters in the model.
- **Answer:** B) It normalizes the inputs of each layer, reducing the impact of poor weight initialization.