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. History of Deep Learning

- **Puzzle:** In the 1980s, a breakthrough in neural network training techniques was developed. This technique improved the efficiency of training deep networks. What is the name of the technique, and which famous scientist is it associated with?
- **Options:**
  - A) Gradient Descent, Geoffrey Hinton
  - B) Backpropagation, Yann LeCun
  - C) Convolution, Geoffrey Hinton
  - D) Recurrent Neural Networks, Yoshua Bengio
- **Answer:** B) Backpropagation, Yann LeCun

2. A Probabilistic Theory of Deep Learning

- **Puzzle:** Imagine you're using a probabilistic model to predict the likelihood of a data point belonging to a class. If your model produces the same output probability regardless of the input, what would this tell you about the effectiveness of your model?
- **Options:**
  - A) The model is overfitting
  - B) The model is underfitting
  - C) The model has learned the optimal parameters
  - D) The model is biased
- **Answer:** B) The model is underfitting

3. Backpropagation and Regularization

- **Puzzle:** If a neural network experiences overfitting, which technique would you use to penalize large weights and prevent overfitting, and how would this affect the network's generalization ability?
- **Options:**
  - A) Data Augmentation; Increases the size of training data
  - B) Dropout; Reduces the complexity of the model
  - C) L2 Regularization; Adds a penalty to large weights
  - D) Batch Normalization; Ensures faster convergence
- **Answer:** C) L2 Regularization; Adds a penalty to large weights

4. Batch Normalization

- **Puzzle:** A neural network experiences slow training progress and unstable gradients. After implementing batch normalization, the network's training improves significantly. What role does batch normalization play in stabilizing the learning process?
- **Options:**
    - A) It increases the learning rate
    - B) It normalizes inputs to each layer to prevent vanishing/exploding gradients
    - C) It increases the number of training epochs
    - D) It reduces the number of layers in the network
- **Answer:** B) It normalizes inputs to each layer to prevent vanishing/exploding gradients

5. VC Dimension and Neural Networks

- **Puzzle:** You have a neural network with a very high VC dimension. What would this imply about the network's ability to classify complex data, and what is the risk associated with a very high VC dimension?
- **Options:**
    - A) The network can generalize well to new data; low risk of overfitting
    - B) The network has limited capacity for learning; high risk of underfitting
    - C) The network has the capacity to model complex data; high risk of overfitting
    - D) The network is simple but very accurate; low risk of overfitting
- **Answer:** C) The network has the capacity to model complex data; high risk of overfitting

6. Deep vs Shallow Networks

- **Puzzle:** If you have two networks: one shallow with fewer layers and one deep with many layers, which one would you expect to have a better capacity to approximate complex functions, and why?
- **Options:**
    - A) The shallow network; because it has fewer parameters
    - B) The deep network; because it can capture hierarchical representations of data
    - C) The shallow network; because it is easier to train
    - D) Both networks would perform equally well
- **Answer:** B) The deep network; because it can capture hierarchical representations of data

7. Convolutional Networks

- **Puzzle:** In a convolutional neural network (CNN), what is the main advantage of using convolutional layers over fully connected layers, especially when dealing with image data?

- **Options:**
  - A) Convolutional layers capture local spatial hierarchies in the data
  - B) Convolutional layers have more parameters than fully connected layers
  - C) Convolutional layers are easier to train than fully connected layers
  - D) Convolutional layers require less computational power than fully connected layers
- **Answer:** A) Convolutional layers capture local spatial hierarchies in the data

8. Generative Adversarial Networks (GAN)

- **Puzzle:** In a GAN, the generator creates fake data, and the discriminator tries to identify whether the data is real or fake. What would happen if the discriminator becomes too powerful compared to the generator?
- **Options:**
  - A) The generator will learn faster
  - B) The generator will fail to improve and stop learning
  - C) The training process will speed up significantly
  - D) The discriminator will no longer be able to distinguish between real and fake data
- **Answer:** B) The generator will fail to improve and stop learning

9. Semi-supervised Learning

- **Puzzle:** You have a semi-supervised learning problem where you only have a small amount of labeled data and a large amount of unlabeled data. Which strategy would you use to make the best use of both types of data?
- **Options:**
  - A) Train only on the labeled data to avoid confusion
  - B) Use unsupervised learning techniques to first pre-train the model, then fine-tune on labeled data
  - C) Discard the unlabeled data and train solely on labeled data
  - D) Use transfer learning from a large pre-trained model
- **Answer:** B) Use unsupervised learning techniques to first pre-train the model, then fine-tune on labeled data

10. Deep vs Shallow Networks (Advanced)

- **Puzzle:** Suppose you have a deep network that trains well on the training data but performs poorly on the test data, while a shallow network performs better on the test data. What could be the reason for this, and how would you address it?
- **Options:**
    - A) The deep network is overfitting; regularization or early stopping can help
    - B) The shallow network is underfitting; increasing its complexity can help
    - C) The deep network is not complex enough; adding more layers can help
    - D) The shallow network is overfitting; regularization can help
- **Answer:** A) The deep network is overfitting; regularization or early stopping can help