

Deep learning and Unsupervised learning

Your Name *

Sihan Wei

What is the **primary** reason for using ReLU instead of a sigmoid activation in deep networks?

1 point

- A. It guarantees better accuracy
- B. It solves vanishing gradients
- C. It has bounded outputs, avoiding exploding activations
- D. It reduces training time by limiting neuron activity

Which of the following are **adaptive learning rate methods**?

1 point

- A. Adagrad
- B. RMSProp
- C. Adam
- D. SGD

What is the purpose of the **dropout** technique during training?

1 point

- A. Speed up training
- B. Prevent overfitting by randomly deactivating neurons
- C. Increase the capacity of the model
- D. Reduce the size of the dataset

Which of the following are true about **L2 regularization** in neural networks?

1 point

- A. It penalizes large weight values to prevent overfitting
- B. It encourages weights to become exactly zero
- C. It modifies the original loss function by adding a term proportional to the squared norm of weights
- D. It is also known as weight decay

Which problem is **most associated with very deep neural networks**?

1 point

- A. Underfitting
- B. Exploding gradients
- C. Vanishing gradients
- D. Over-training

Which of the following techniques are used to **mitigate overfitting**?

1 point

- A. Dropout
- B. Adaptive learning rate
- C. L2 regularization
- D. Early stopping

Which of the following statements correctly describe **generative models**?

1 point

- A. They model the decision boundary between classes
- B. They can generate new data samples from the learned distribution
- C. They model the joint distribution $p(x, y)$ or the marginal $p(x)$
- D. Logistic regression is an example of a generative model

Which of the following are **limitations of K-means**?

1 point

- A. Requires labeled data
- B. Sensitive to initialization
- C. Must specify k in advance
- D. Finds global optima

Which of the following models **can be used to generate new samples** after training?

1 point

- A. Logistic Regression
- B. K-means
- C. GAN
- D. Lasso Regression

Which of the following are **assumptions or limitations** of the K-Means clustering algorithm?

1 point

- A. K-Means minimizes the sum of squared distances between points and their assigned cluster centroids
- B. K-Means is guaranteed to find the global optimum of the objective function
- C. K-Means is sensitive to outliers
- D. K-Means assigns each point to exactly one cluster (hard clustering)

This content is neither created nor endorsed by Google.

Google Forms