

Quiz & Solutions

Multiple-choice questions.

1. What is data augmentation?

- A) Creating modified copies of existing data
 - B) A technique to reduce overfitting
 - C) A method that causes overfitting
 - D) A technique to generate synthetic training data
 - E) None of the above
-

2. Why are data important for advanced machine learning applications?

- A) It improves model performance
 - B) It directly affects model performance
 - C) It is essential for hyperparameter tuning
 - D) It reduces computational overhead
 - E) It is required for both training and evaluating models
-

3. Why use data augmentation?

- A) To reduce dataset size
 - B) To lower data diversity
 - C) To increase computational cost
 - D) To improve model performance while minimizing the need for new data collection
 - E) To intentionally reduce accuracy
-

4. What are the main purposes of data augmentation?

- A) Increasing dataset size
 - B) Addressing class imbalance
 - C) Reducing model complexity
 - D) Both A and C
 - E) None of the above
-

5. What is a typical data augmentation workflow?

- A) Evaluate model performance → choose augmentation technique → apply it → re-evaluate
 - B) Measure performance → augment data → shuffle dataset
 - C) Measure performance → select technique → augment data
 - D) Augment data → evaluate performance → select technique
 - E) Shuffle data → measure performance → augment data
-

6. At which stage is data augmentation usually applied?

- A) Validation
 - B) Test
 - C) Data preprocessing
 - D) Training
 - E) Hyperparameter optimization
-

7. What is a major benefit of data augmentation?

- A) Increased computational cost
 - B) Higher risk of noise
 - C) Improved model generalization
 - D) Limited applicability
 - E) Improved fairness
-

8. What is a common drawback of data augmentation?

- A) Improved generalization
 - B) Cost-efficiency
 - C) Potential to introduce noise or artifacts
 - D) Better fairness
 - E) Prevention of overfitting
-

9. What is a risk associated with data augmentation?

- A) Unintended bias or fairness issues
 - B) Increased computational requirements
 - C) Injection of noisy or unrealistic data
 - D) All of the above
 - E) None of the above
-

10. What is a key ethical concern in data augmentation?

- A) Increased hardware demands
- B) Risk of noise
- C) Lack of impact in some domains
- D) Potential bias and fairness issues
- E) Enhanced generalization