

# Machine Learning (ML) MCQs

**1. Application of machine learning methods to large databases is called**

- A. data mining.
- B. artificial intelligence
- C. big data computing
- D. internet of things

Answer: A

---

**2. If machine learning model output involves target variable then that model is called as**

- A. descriptive model
- B. predictive model
- C. reinforcement learning
- D. all of the above

Answer: B

---

**3. In what type of learning labelled training data is used**

- A. unsupervised learning
- B. supervised learning
- C. reinforcement learning
- D. active learning

Answer: B

---

**4. In following type of feature selection method we start with empty feature set**

- A. forward feature selection
- B. backward feature selection
- C. both a and b??
- D. none of the above

Answer: A

---

**5. In PCA the number of input dimensiona are equal to principal components**

- A. true
- B. false

Answer: A

---

**6. PCA can be used for projecting and visualizing data in lower dimensions.**

- A. true
- B. false

Answer: A

---

**7. Which of the following is the best machine learning method?**

- A. scalable
- B. accuracy
- C. fast
- D. all of the above

Answer: D

---

**8. What characterize unlabeled examples in machine learning**

- A. there is no prior knowledge
- B. there is no confusing knowledge
- C. there is prior knowledge
- D. there is plenty of confusing knowledge

Answer: D

---

**9. What does dimensionality reduction reduce?**

- A. stochastics
- B. collinerity
- C. performance
- D. entropy

Answer: B

---

**10. Data used to build a data mining model.**

- A. training data
- B. validation data
- C. test data
- D. hidden data

Answer: A

---

**11. The problem of finding hidden structure in unlabeled data is called...**

- A. supervised learning
- B. unsupervised learning
- C. reinforcement learning

D. none of the above

Answer: B

---

**12. Of the Following Examples, Which would you address using an supervised learning Algorithm?**

A. given email labeled as spam or not spam, learn a spam filter

B. given a set of news articles found on the web, group them into set of articles about the same story.

C. given a database of customer data, automatically discover market segments and group customers into different market segments.

D. find the patterns in market basket analysis

Answer: A

---

**13. Dimensionality Reduction Algorithms are one of the possible ways to reduce the computation time required to build a model**

A. true

B. false

Answer: A

---

**14. You are given reviews of few netflix series marked as positive, negative and neutral. Classifying reviews of a new netflix series is an example of**

A. supervised learning

B. unsupervised learning

C. semisupervised learning

D. reinforcement learning

Answer: A

---

**15. Which of the following is a good test dataset characteristic?**

A. large enough to yield meaningful results

B. is representative of the dataset as a whole

C. both a and b

D. none of the above

Answer: C

---

**16. Following are the types of supervised learning**

A. classification

B. regression

- C. subgroup discovery
- D. all of the above

Answer: D

---

**17. Type of matrix decomposition model is**

- A. descriptive model
- B. predictive model
- C. logical model
- D. none of the above

Answer: A

---

**18. Following is powerful distance metrics used by Geometric model**

- A. euclidean distance
- B. manhattan distance
- C. both a and b??
- D. square distance

Answer: C

---

**19. The output of training process in machine learning is**

- A. machine learning model
- B. machine learning algorithm
- C. null
- D. accuracy

Answer: A

---

**20. A feature F1 can take certain value: A, B, C, D, E, & F and represents grade of students from a college. Here feature type is**

- A. nominal
- B. ordinal
- C. categorical
- D. boolean

Answer: B

---

**21. PCA is**

- A. forward feature selection
- B. backward feature selection

- C. feature extraction
- D. all of the above

Answer: C

---

**22. Dimensionality reduction algorithms are one of the possible ways to reduce the computation time required to build a model.**

- A. true
- B. false

Answer: A

---

**23. Which of the following techniques would perform better for reducing dimensions of a data set?**

- A. removing columns which have too many missing values
- B. removing columns which have high variance in data
- C. removing columns with dissimilar data trends
- D. none of these

Answer: A

---

**24. Supervised learning and unsupervised clustering both require which is correct according to the statement.**

- A. output attribute.
- B. hidden attribute.
- C. input attribute.
- D. categorical attribute

Answer: C

---

**25. What characterize is hyperplane in geometrical model of machine learning?**

- A. a plane with 1 dimensional fewer than number of input attributes
- B. a plane with 2 dimensional fewer than number of input attributes
- C. a plane with 1 dimensional more than number of input attributes
- D. a plane with 2 dimensional more than number of input attributes

Answer: B

---