More	narendran.ss@gmail.com Dashboard Sign Out

Coding Tech Life

Machine Learning Quiz 1 BITS PILANI WILP

Machine Learning (ISZC464) Quiz 1
BITS PILANI WILP - 2017

1. Averaging the output of multiple decision trees helps
Select one:
a. Increase bias
b. Increase variance
c. Decrease bias
d. Decrease variance

Ans: d. Decrease variance

2. Given genetic (DNA) data from a person, predict the odds of him/her developing diabetes over the next 10 years. What kind of learning problem is this?
Select one:
a. None of the given answers
b. Unsupervised Learning
c. Supervised Learning
d. Reinforcement Learning

Ans: c. Supervised Learning

3. Given a large dataset of medical records from patients suffering from heart disease, try to learn whether there might be different clusters of such patients for which we might tailor separate treatments. What kind of learning problem is this?
Select one:
a. Supervised Learning
b. Unsupervised Learning
c. None of the given answers
d. Reinforcement Learning

Ans: b. Unsupervised Learning

4. In farming, given data on crop yields over the last 50 years, learn to predict next year's crop yields. What kind of learning problem is this?
Select one:
a. None of the given answers
b. Unsupervised Learning
c. Reinforcement Learning
d. Supervised Learning

Ans: d. Supervised Learning

5. Suppose we wish to calculate $P(H \mid E1, E2)$ and we have no conditional independence information. Which of the following sets are sufficient for computing this (minimal set)?
Select one:
a. $P(E1, E2 \mid H)$, $P(H)$, $P(E1 \mid H)$, $P(E2 \mid H)$
b. $P(H)$, $P(E1 \mid H)$, $P(E2 \mid H)$
c. $P(E1, E2)$, $P(H)$, $P(E1 \mid H)$, $P(E2 \mid H)$
d. $P(E1, E2)$, $P(H)$, $P(E1, E2 \mid H)$

Ans: c. $P(E1, E2)$, $P(H)$, $P(E1 \mid H)$, $P(E2 \mid H)$

6. Which of the following strategies cannot help reduce overfitting in decision trees?
Select one:
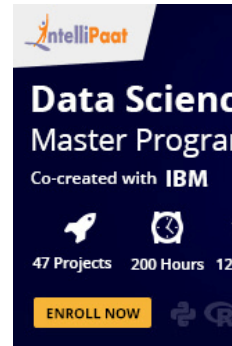a. Enforce a maximum depth for the tree
b. Enforce a minimum number of samples in leaf nodes
c. Make sure each leaf node is one pure class
d. Pruning

Ans: c. Make sure each leaf node is one pure class

7. Suppose we wish to calculate $P(H \mid E1, E2)$ and we know that $P(E1 \mid H, E2) = P(E1 \mid H)$ for all the values of H, E1, E2. Now which of the following sets are sufficient?
Select one:
a. $P(E1, E2 \mid H)$, $P(H)$, $P(E1 \mid H)$, $P(E2 \mid H)$

Search This Blog

Search

Blog Archive

Labels

Advanced Data Mining Android AngularJs Artificial Intelligence Automation BITS Comprehensive Test BITS Mid-Semester Test BITS Previous Question Paper BITS Quiz BITS-WILP Bootstrap ClassDiagram Cloud Computing Cooking CSS d3 Data Mining Data Storage Technologies and Networks Data Structures and Algorithms Design Data

b. P(E1, E2) , P(H), P(E1|H), P(E2|H)
c. P(E1, E2), P(H), P(E1, E2| H)
d. P(H), P(E1| H), P(E2|H)

Ans: b. P(E1, E2) , P(H), P(E1|H), P(E2|H)

8. Take a collection of 1000 essays written on the US Economy, and find a way to automatically group these essays into a small number of groups of essays that are somehow "similar" or "related". What kind of learning problem is this?
Select one:
a. None of the given answers
b. Reinforcement Learning
c. Unsupervised Learning
d. Supervised Learning

Ans: c. Unsupervised Learning

9. Suppose you are working on weather prediction, and you would like to predict whether or not it will be raining at 5pm tomorrow. You want to use a learning algorithm for this. What machine learning task is this?
Select one:
a. Clustering
b. Classification
c. None of the given answers
d. Regression

Ans: b. Classification

10. You've just finished training a decision tree for spam classification, and it is getting abnormally bad performance on both your training and test sets. You know that your implementation has no bugs, so what could be causing the problem?
Select one:
a. You need to increase the learning rate.
b. Your decision trees are too shallow.
c. All of the other given options.
d. You are overfitting.

Ans: a. You need to increase the learning rate.

11. Which of the following statements about classification is true?
Select one:
a. As the number of data points grows to infinity, the MAP estimate approaches the MLE estimate for all possible priors. In other words, given enough data, the choice of prior is irrelevant
b. No classifier can do better than a naive Bayes classifier if the distribution of the data is known
c. Density estimation (using say, the kernel density estimator) can be used to perform classification
d. The depth of a learned decision tree can be larger than the number of training examples used to create the tree

Ans: c. Density estimation (using say, the kernel density estimator) can be used to perform classification

12. Consider the task of examining a large collection of emails that are known to be spam email, to discover if there are sub-types of spam mail. What kind of learning problem is this?
Select one:
a. Reinforcement Learning
b. Supervised Learning
c. Unsupervised Learning
d. None of the given answers

Ans: c. Unsupervised Learning

13. Suppose you are working on stock market prediction, and you would like to predict the price of a particular stock tomorrow (measured in dollars). You want to use a learning algorithm for this. What machine learning task is this?
Select one:
a. Classification
b. Clustering
c. Regression
d. None of the given answers

Ans: c. Regression

14. For polynomial regression, which one of these structural assumptions is the one that most affects the trade-off between underfitting and overfitting
Select one:
a. The assumed variance of the Gaussian noise
b. The use of a constant-term unit input
c. Whether we learn the weights by gradient descent
d. The polynomial degree

Ans: d. The polynomial degree

15. Which of the following statements is true?

Select one:

a. Given m data points, the training error converges to the true error as m →∞

b. Decision tree is learned by minimizing information gain

c. Linear regression estimator has the smallest variance among all unbiased estimators

d. A classifier trained on less training data is less likely to overfit

Previous page

Ans: a. Given m data points, the training error converges to the true error as m →∞

16. Given 50 articles written by male authors, and 50 articles written by female authors, learn to predict the gender of a new manuscript's author (when the identity of this author is unknown). What kind of learning problem is this?

Select one:

a. None of the given answers

b. Supervised Learning

c. Reinforcement Learning

d. Unsupervised Learning

Ans: b. Supervised Learning

Labels: BITS Quiz, BITS-WILP, Machine Learning, Machine Learning Quiz

1 comment:

akhilapriya404 November 15, 2018 at 1:47 AM

Thanks you for sharing the article. The data that you provided in the blog is infromative and effectve. Through you blog I gained so much knowledge. Also check my collection at machine learning online training Blog

Reply

```
Enter your comment...
```

Comment as: Narendran (Go ▼)   Sign out

Publish   Preview   ☐ Notify me

Newer Post   Home   Older Post

Subscribe to: Post Comments (Atom)

Simple theme. Powered by Blogger.