Print

Quiz: Introduction to Predictive Modeling: Predictive Modeling and Decision Trees

Your Score: 100%

Congratulations! Your score indicates that you have mastered the topics in this lesson. You can review the feedback and when you're finished, exit the lesson.

1. Which type of data set is used to monitor and tune a predictive model?

   a. Training
   b. Validation
   c. Testing
   d. Score

Your answer: b
Correct answer: b

A validation data set is used for monitoring and tuning a model to improve its generalization. The tuning process usually involves selecting among models of different types and complexities. The tuning process optimizes the selected model on the validation data.

---

2. Which of the following is an essential task for any predictive model?

   a. predict new cases
   b. select useful inputs
   c. optimize complexity
   d. all of the above

Your answer: d
Correct answer: d

Predictive models are widely used and come in many varieties. Any model must perform three essential tasks: predict new cases, select useful inputs, and optimize complexity. Different modeling tools use different methods to complete each task.

---

3. Which type of prediction orders cases based on the inputs' relationship to the target?

   a. Decision
   b. Classification
   c. Ranking
   d. Estimate

Your answer: c
Correct answer: c

Ranking predictions order cases based on the inputs' relationships with the target. Using the training data, the prediction model attempts to rank high value cases higher than low value cases. It is assumed that a similar pattern exists in the scoring data so that high value cases have high scores. The actual scores produced are inconsequential; only the relative order matters. A common example of a ranking prediction is a credit score.

4. When a decision tree grows, the best split for an input is a split that yields what?

   a. a maximal tree
   b. a contingency table
   c. a depth adjustment
   d. the highest logworth

Your answer: d
Correct answer: d

The best split for an input is the split that yields the highest logworth. This value corresponds to the most statistically significant Chi-square test for independence.

5. Decision Tree models use pruning to adjust model complexity and avoid the potential problem known as what?

   a. overfitting
   b. accuracy
   c. concordance
   d. misclassification

Your answer: a
Correct answer: a

The maximal tree represents the most complicated model you are willing to construct from a set of training data. To avoid potential overfitting, many predictive modeling procedures offer some mechanism for adjusting model complexity. For decision trees, this process is known as pruning.

Close