Oracle Database 19c: Data Mining Techniques LVC

Oracle Database

DURATION	MODULES	COURSE CODE
2 Days	**9 Lectures**	**—**

Course Overview

The course covers the following topics: Basic concepts and primary terminology of data mining Steps associated with the data mining process Use of Oracle Data Miner 19c to perform data mining Options for deploying data mining predictive results

What You Will Learn

- I Introduction
- Course Goals I-2
- Course Agenda I-3
- Summary I-4

Data Mining Concepts and Terminology

- Objectives
- What is data mining?
- Who uses data mining?
- Defining Key Data Mining Properties
- Data Mining Categories
- Examining Supervised Data Mining
- Supervised Data Mining Steps
- Supervised Data Mining Techniques
- Supervised Data Mining Algorithms
- Examining Unsupervised Data Mining
- Unsupervised Data Mining Techniques
- Unsupervised Data Mining Algorithms
- What Data Mining Can and Cannot Do
- Summary

The Data Mining Process

- Objectives

- Steps in the Data Mining Process
- Step 1: Defining the Problem
- Step 2: Selecting, Preparing, and Preprocessing Data
- Data Preparation Tips
- Data Preparation: Aggregation
- Data Preparation: Mining Attribute Types
- Automatic Data Preparation (ADP)
- Step 3: Building and Evaluating Models
- Model-Building Tasks
- Training and Testing Models: Supervised Learning
- Evaluating Model Results
- Applying the Selected Models
- Step 4: Deploying Results
- Summary

Introducing Oracle Data Miner

- Objectives
- Oracle Data Mining: Architecture
- Oracle Data Mining: Related Uses
- Oracle Data Miner
- Setting Up Oracle Data Miner (Using SQL Developer)
- Installing SQL Developer
- Setting Up Data Miner
- The Data Miner Repository Installation Process
- Creating a Data Miner Project and Workflow
- Introducing the Data Miner Interface
- Examining Oracle Data Miner Nodes
- Previewing a Data Miner Workflow
- Summary

Using Classification Models

- Objectives
- Revisiting Supervised Learning
- Supervised Learning: Examining Classification Techniques
- Using Classification Models
- Building a Workflow
- The Starting Point for Any Workflow: A Data Source
- Adding a Data Source to the Workflow
- Defining a Data Source: Step 1
- Viewing Source Data
- Defining a Data Source: Step 2
- Completed Data Source Node
- Previewing the Source Data
- Adding an Explore Data Node
- Connecting Workflow Nodes

- Specifying a “Group By” for the Explore Data Node
- Running the Explore Data Node
- Viewing the Explore Data Results
- Examining the Results
- Creating Classification Models
- Adding a Classification Node
- Classification Node Default Properties
- Connecting Data Source and Class Build Nodes
- Selecting Target and Case ID Attributes
- Viewing Advanced Settings
- Accessing Help
- Specifying Advanced Settings
- Specifying Test Properties
- Building the Models
- Successful Build Indicators
- Viewing the Event Log
- Comparing the Models
- Comparing Test Results: Performance
- Comparing Test Results: Lift
- Comparing Test Results: Performance Matrix
- Comparing Test Results: Profit
- Tuning a Model
- Selecting a Specific Model
- Navigating the Decision Tree Model Interface
- Examining the Decision Tree Test Results
- Viewing the Leaf Node in the Decision Tree
- Scoring: Applying the Model to Your Data
- Step 1: Specifying the Desired Model
- Step 2: Adding a New Data Source Node
- Modifying a Node Name
- Step 3: Adding an Apply Node
- Step 4: Connecting the Nodes
- Step 5: Creating the Output Format
- Step 6: Running the Apply Node
- Viewing Predictive Results
- Sorting Results
- Examining Predictions
- Summary

Using Regression Models

- Objectives
- Supervised Learning: Examining Regression Techniques
- Using Regression Models
- Creating a New Workflow
- Adding the Data Source

- Understanding Target and Input Variables
- Adding, Connecting, and Running an Explore Data Node
- Exploring Source Data Statistics
- Identifying Non-Input Variables
- Filtering the Source Data
- Deselecting Columns
- Running the Source Data Filter
- Creating Regression Models
- Adding a Regression Node
- Regression Node Default Properties
- Connecting and Editing the Regression Node
- Editing Advanced Settings: GLM Regression
- Viewing Advanced Settings: SVM Regression
- Building the Models
- Successful Build Indicators
- Viewing Models and Results
- Viewing GLM Model Details
- Viewing GLM Model Settings
- Viewing GLM Model Coefficients
- Viewing SVM Model Information
- Sorting Model Coefficients
- Viewing Test Results for the GLM Model
- GLM Model Test Results: Residual Tab
- GLM Model Test Results: Predicted and Actual
- Viewing Test Results for the SVM Model
- SVM Model Test Results: Predicted and Actual
- Comparing Test Results
- Comparing GLM and SVM Test Results
- Scoring Data with the Selected Model
- Step 1: Specifying the Chosen Model
- Step 2: Adding a New Data Source Node
- Removing Non-Input and Target Variables
- Modifying the Node Name
- Step 3: Adding an Apply Node
- Step 4: Connecting the Nodes
- Step 5a: Modifying the Output Format
- Step 5b: Modifying the Output Format
- Step 5c: Modifying the Output Format
- Step 6: Running the Apply Node
- Viewing Predictive Results
- Sorting Results
- Examining Predictions
- Summary

Using Clustering Models

- Objectives
- Revisiting Unsupervised Data Mining
- Unsupervised Data Mining: Clustering
- Using a Clustering Model
- Creating the New Workflow
- Using Existing Workflow Objects
- Connecting and Editing the Explore Data Node
- Specifying a “Group By” for the Explore Data Node
- Running and Viewing the Explore Data Node
- Exploring Source Data Statistics
- Creating Clustering Models
- Adding a Clustering Node
- Default Properties of Clustering Nodes
- Connecting the Nodes and Modifying Node Properties
- Examining Cluster Model Settings
- K-Means Advanced Algorithm Settings
- O-Cluster Advanced Algorithm Settings
- Expectation Maximization Advanced Algorithm Settings
- Building the Models
- Viewing the Models
- Viewing K-Means Model Results
- Navigating the K-Means Tree
- Viewing Results with Cluster Rule
- Viewing Other Cluster Definitions
- Viewing Cluster Details
- Comparing Clusters
- Multiple Row Selection with Cluster Compare
- Renaming Clusters
- Viewing the O-Cluster Model Results
- Viewing the Expectation Maximization Model Results
- Applying a Clustering Model
- Creating an Output Format: Apply Columns
- Creating an Output Format: Data Columns
- Running the Model and Viewing Data
- Viewing Results
- Viewing Sorted Results
- Summary

Performing Market Basket Analysis

- Objectives
- Unsupervised Data Mining: Association Rules
- Using an Association Rules Model
- Creating a New Project and Workflow

- Adding a Data Source from a New Schema
- Selecting a New Table or View
- Previewing the Data Source
- Examining Sorted Data
- Creating an Association Rules Model
- Adding an Association Node
- Connecting Data Source and Assoc Build Nodes
- Specifying Required Input Items
- Specifying Advanced Settings
- Building the Model
- Viewing the Model
- Viewing Model Results
- Examining Model Rules
- Creating Model Rule Filters
- Applying Model Rule Filters
- Viewing Model Itemsets
- Creating Model Details Output
- Viewing Model Details
- Using Descriptive Columns in the Output
- Adding a Join Node
- Defining the Join
- Specifying Output Columns for the Join
- Defining a New Assoc Build Node
- Viewing Model Results
- Viewing Descriptive Model Rules
- Creating New Model Details Output
- Viewing Descriptive Model Details
- Summary

Performing Anomaly Detection

- Objectives
- Unsupervised Data Mining: Anomaly Detection
- Using an Anomaly Detection Model
- Creating the New Workflow
- Adding a Data Source
- Browsing Source Data Attributes
- Previewing the Source Data
- Creating an Anomaly Detection Model
- Adding an Anomaly Detection Node
- Connecting the Nodes and Selecting the Case ID
- Specifying the SVM Algorithm Kernel Function
- Building the Model
- Viewing Model Results
- Viewing “Normal” Results
- Viewing “Anomalous” Results

- Comparing Model Results
- Applying the Model
- Creating an Output Format: Apply Columns
- Creating an Output Format: Data Columns
- Running the Model and Viewing Data
- Sorting Output Results
- Examining Results
- Summary

Deploying Data Mining Results

- Objectives
- Agenda
- Creating a Table or View of Results
- Modifying the Output Format
- Generating the Output and Viewing Results
- Viewing Output Results
- Using Output Results
- Agenda
- SQL Code Automatically Generated by Data Miner
- Deploying SQL Code to the Clipboard
- Accessing Deployed Code
- Executing Deployed Code
- Modifying Deployed Code
- Summary