Oracle® Data Mining

Administrator's Guide 10*g* Release 2 (10.2) **B14338-02**

November 2005



Oracle Data Mining Administrator's Guide, 10g Release 2 (10.2)

B14338-02

Copyright © 2005, Oracle. All rights reserved.

The Programs (which include both the software and documentation) contain proprietary information; they are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright, patent, and other intellectual and industrial property laws. Reverse engineering, disassembly, or decompilation of the Programs, except to the extent required to obtain interoperability with other independently created software or as specified by law, is prohibited.

The information contained in this document is subject to change without notice. If you find any problems in the documentation, please report them to us in writing. This document is not warranted to be error-free. Except as may be expressly permitted in your license agreement for these Programs, no part of these Programs may be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose.

If the Programs are delivered to the United States Government or anyone licensing or using the Programs on behalf of the United States Government, the following notice is applicable:

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, use, duplication, disclosure, modification, and adaptation of the Programs, including documentation and technical data, shall be subject to the licensing restrictions set forth in the applicable Oracle license agreement, and, to the extent applicable, the additional rights set forth in FAR 52.227-19, Commercial Computer Software—Restricted Rights (June 1987). Oracle Corporation, 500 Oracle Parkway, Redwood City, CA 94065

The Programs are not intended for use in any nuclear, aviation, mass transit, medical, or other inherently dangerous applications. It shall be the licensee's responsibility to take all appropriate fail-safe, backup, redundancy and other measures to ensure the safe use of such applications if the Programs are used for such purposes, and we disclaim liability for any damages caused by such use of the Programs.

Oracle, JD Edwards, PeopleSoft, and Retek are registered trademarks of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.

The Programs may provide links to Web sites and access to content, products, and services from third parties. Oracle is not responsible for the availability of, or any content provided on, third-party Web sites. You bear all risks associated with the use of such content. If you choose to purchase any products or services from a third party, the relationship is directly between you and the third party. Oracle is not responsible for: (a) the quality of third-party products or services; or (b) fulfilling any of the terms of the agreement with the third party, including delivery of products or services and warranty obligations related to purchased products or services. Oracle is not responsible for any loss or damage of any sort that you may incur from dealing with any third party.

Contents

eface	. Vi
Audience	vi
Documentation Accessibility	vi
Related Documents	vii
Conventions	vii
Installing Oracle Database for Data Mining	
Who Needs to Install Oracle Database?	1-1
Understanding Oracle Universal Installer	1-2
Installation Directories	1-2
Installation Methods	1-2
Installation Types	1-3
Creating a Database During Installation	
Preparing for Installation	1-3
Hardware Requirements	1-3
Software Requirements	1-4
Preexisting Oracle Products	1-4
Dynamic IP	1-4
Installing the Data Mining Option	1-4
Adding the Data Mining Option to a Database	1-9
Removing the Data Mining Option	1-10
Installing the Data Mining Scoring Engine	1-11
Performing a Custom Installation	1-11
Adding the Scoring Engine to an Existing Installation	1-13
Removing the Scoring Engine	1-13
Verifying Installation of Data Mining Components	1-14
Tips for Linux Installations	1-15
Upgrading from Oracle Database 10g Release 1 (10.1)	1-15
Post-Upgrade Steps for Data Mining	1-15
Upgrading PL/SQL Models	1-16
Downgrading Oracle Data Mining	1-17
Administering Oracle Database for Data Mining	
Local Administration on Microsoft Windows	2-1
Oracle Enterprise Manager Database Control	2-1
	Documentation Accessibility Related Documents Conventions Installing Oracle Database for Data Mining Who Needs to Install Oracle Database? Understanding Oracle Universal Installer Installation Directories Installation Methods Installation Types. Creating a Database During Installation Preparing for Installation Hardware Requirements Software Requirements Preexisting Oracle Products. Dynamic IP Installing the Data Mining Option Adding the Data Mining Option to a Database Removing the Data Mining Option Installing the Data Mining Scoring Engine Performing a Custom Installation Adding the Scoring Engine to an Existing Installation Removing the Scoring Engine to an Existing Installation Removing the Scoring Engine to Data Mining Components Tips for Linux Installations Upgrading from Oracle Database 10g Release 1 (10.1) Post-Upgrade Steps for Data Mining Upgrading PL/SQL Models Downgrading Oracle Database for Data Mining Local Administering Oracle Database for Data Mining Local Administration on Microsoft Windows

	SQL*Plus	2-2
	Database Configuration Assistant	2-3
	Oracle Universal Installer	2-3
	Oracle Services	2-3
	Local Administration on Linux	2-3
	Remote Administration	2-4
	Creating Data Mining Users	2-4
	Creating Tablespaces	2-4
	Creating Users	2-7
	About the DMSYS Schema	2-11
	Exporting and Importing Data Mining Models	2-11
	Prerequisites	2-12
	PL/SQL APIs for Exporting and Importing Models	2-13
	Java APIs for Exporting and Importing Models	2-13
	Tables Created By Exporting and Importing Models	
	Example: Exporting and Importing Models	
3	Installing Tools for Data Mining	
	Oracle Data Miner	3-1
	Installing Oracle Data Miner on Windows	3-1
	Installing Oracle Data Miner on Linux	3-1
	Connecting Oracle Data Miner to Oracle Database	. 3-2
	Oracle Spreadsheet Add-In for Predictive Analytics	3-3
	Installation Steps	3-3
	Adding the Spreadsheet Add-In to Excel	3-4
	Installing Oracle Client	3-4
	Creating an Oracle Net Service Name	. 3-6
	Connecting Excel to Oracle Database	3-9
4	Installing and Using the Demo Programs	
-	Requirements Checklist	1 1
	Obtaining the Demo Programs	
	Downloading from OTN	
	Installing Oracle Database Companion	
	Verifying the Installation	
	Creating a Demo User	
	Using SQL*Plus to Create the User	
	Using Enterprise Manager to Create the User	
	Granting Access Rights	
	Populating the Schema	
	A Sample User Scenario	
	Examining the Data	
	Customer Data for Data Mining	
	Market Basket Data for Association Rules	4-0 4-10
	Customer Data for Text Mining	4-10
	PL/SQL Demo Programs	4-10
	PL/SOL Program Summaries	4-11
	1 L/ UVL 1 1VE1011 DUITHIOTUS	T 14

Data Mining SQL Functions	4-13
Running the PL/SQL Programs	4-14
Java Demo Programs	4-14
Java Program Summaries	4-15
Preparing to Run the Java Programs	4-16
Running the Java Programs	4-17
Text Mining Demo Programs	4-17
Text Mining in PL/SQL	4-18
Text Mining in Java	4-19
BLAST Demo Program	4-20
Preparing to Run the BLAST Demo	4-20
Running the BLAST Table Functions	4-20

Index

Preface

The *Oracle Data Mining Administrator's Guide* explains how to install the various components of Oracle Data Mining and perform basic administration tasks.

The preface contains these topics:

- Audience
- Documentation Accessibility
- Related Documents
- Conventions

Audience

This guide is intended for a spectrum of users who need to install the software that supports Oracle Data Mining:

- Individual users who want to evaluate Oracle Data Mining
- Departmental DBAs who manage an instance of Oracle Database for Data Mining users
- Enterprise DBAs and IT professionals who need to support Data Mining users as part of administering Oracle Database for a larger user community

Note: Although this manual provides DBAs with the basic information they need to install and administer Oracle Data Mining, this manual is targeted primarily at the individual user.

Documentation Accessibility

Our goal is to make Oracle products, services, and supporting documentation accessible, with good usability, to the disabled community. To that end, our documentation includes features that make information available to users of assistive technology. This documentation is available in HTML format, and contains markup to facilitate access by the disabled community. Accessibility standards will continue to evolve over time, and Oracle is actively engaged with other market-leading technology vendors to address technical obstacles so that our documentation can be accessible to all of our customers. For more information, visit the Oracle Accessibility Program Web site at

http://www.oracle.com/accessibility/

Accessibility of Code Examples in Documentation

Screen readers may not always correctly read the code examples in this document. The conventions for writing code require that closing braces should appear on an otherwise empty line; however, some screen readers may not always read a line of text that consists solely of a bracket or brace.

Accessibility of Links to External Web Sites in Documentation

This documentation may contain links to Web sites of other companies or organizations that Oracle does not own or control. Oracle neither evaluates nor makes any representations regarding the accessibility of these Web sites.

TTY Access to Oracle Support Services

Oracle provides dedicated Text Telephone (TTY) access to Oracle Support Services within the United States of America 24 hours a day, seven days a week. For TTY support, call 800.446.2398.

Related Documents

This manual supplements the information you will find in other Oracle Database 10*g* manuals, especially:

- The *Oracle Database Installation Guide* for your platform.
- *Oracle Database 2 Day DBA*
- Oracle Database Administrator's Guide

For more information about Oracle Data Mining, see the following Oracle Database 10*g* manuals:

- Oracle Data Mining Concepts
- Oracle Data Mining Java API Reference
- Oracle Data Mining Application Developer's Guide
- Oracle Database PL/SQL Packages and Types Reference (Search for "Data Mining")
- Oracle Database SQL Reference (Search for "Data Mining")

Conventions

The following text conventions are used in this document:

Convention	Meaning
boldface	Boldface type indicates graphical user interface elements associated with an action, or terms defined in text or the glossary.
italic	Italic type indicates book titles, emphasis, or placeholder variables for which you supply particular values.
monospace	Monospace type indicates commands within a paragraph, URLs, code in examples, text that appears on the screen, or text that you enter.

Installing Oracle Database for Data Mining

This chapter explains how to install Oracle Database to support Data Mining applications and development. It is directed primarily at users who want to install Oracle Database on a personal computer or laptop for their own use.

This chapter contains the following topics:

- Who Needs to Install Oracle Database?
- Understanding Oracle Universal Installer
- Preparing for Installation
- Installing the Data Mining Option
- Installing the Data Mining Scoring Engine
- Verifying Installation of Data Mining Components
- Tips for Linux Installations
- Upgrading from Oracle Database 10g Release 1 (10.1)

Note: You should use this chapter along with your Oracle Database installation documentation. This chapter does not attempt to provide comprehensive installation instructions for Oracle Database.

Who Needs to Install Oracle Database?

If you want to use Oracle Data Mining, then you must have access to Oracle Database. You may be able to access an existing installation instead of installing Oracle Database yourself. Contact your Oracle database administrator and inquire whether you can access an instance of Oracle Database 10g Release 2 that is operating in 10.1 or later compatibility mode.

If you can access an existing installation, then ask your Oracle DBA to do the following:

- Install the Sample Schemas, if they are not already installed.
- Install the Data Mining demo programs, as described in "Installing Oracle Database Companion" on page 4-2.
- 3. Create a user name. See "Creating Data Mining Users" on page 2-4 and "Creating a Demo User" on page 4-4.
- **4.** Provide you with the connection information.

Meanwhile, you can install a Data Mining client on your personal computer, as described in Chapter 3.

If you do not have access to an existing installation, then follow the instructions provided in this chapter for installing Oracle Database.

Understanding Oracle Universal Installer

To install Oracle Database, you use Oracle Universal Installer. This installation program is nearly the same on all platforms, with differences that arise because of the differences in the operating systems. See "Oracle Universal Installer" on page 2-3

This chapter provides step-by-step instructions for using Oracle Universal Installer to install Oracle Database on a Microsoft Windows platform. If you use a different platform, you can still derive the preferred settings for Data Mining from these instructions. However, you should be careful to perform any pre- and post-installation steps specific to your platform, because they may be critical to a successful installation. Refer to the *Installation Guide* for your platform.

See Also:

- Oracle Database Quick Installation Guide for Microsoft Windows (32-Bit)
- Oracle Database Installation Guide for Microsoft Windows
- *Oracle Database 2 Day DBA*

Installation Directories

There are four sets of installation media for Oracle Database. If you download the software instead of installing from a physical media pack, you will create four separate installation directories on your computer.

- **Database**: Installation directory for Oracle Database 10g Release 2 (10.2)
- **Companion**: Installation directory for demo program files. See Chapter 4.
- Client: Installation directory for client tools. Required by Oracle Spreadsheet Add-In for Predictive Analytics (See Chapter 3).
- **Doc**: Installation directory for documentation

Installation Methods

Oracle Universal Installer provides two methods for installing the Oracle Database software:

- Basic: This installation method can be accomplished quickly and requires minimal user input. It installs the software and optionally creates a general-purpose database. It is the default installation method.
- Advanced: This installation method offers a number of choices and may take longer to complete. It installs the specified software components and optionally creates a database. You can choose a general-purpose database, or you can specify that the database be optimized for data warehousing or transaction processing.

You can use either the Basic or the Advanced method to install Oracle Database with support for Data Mining. This chapter provides instructions for performing a Basic installation.

To install the Data Mining Scoring Engine, you must use the Advanced method.

Installation Types

Whether you use the Basic or Advanced installation method, you can specify one of the following installation types for Oracle Database:

- Enterprise Edition: Installs licensable Oracle Database options, including Data Mining. Includes configuration and management tools in addition to all of the products that are installed during a Standard Edition installation.
- **Standard Edition**: Installs an integrated set of management tools, web features, and facilities for creating business applications.
- Personal Edition: Installs Enterprise Edition software, but supports only a single user development and deployment environment.
- **Custom**: Installs the individual components that you specify.

Data Mining is only available in the Enterprise Edition. See "Installing the Data Mining Option" on page 1-4.

To install the Data Mining Scoring Engine, you must perform a Custom installation. See "Installing the Data Mining Scoring Engine" on page 1-11.

Creating a Database During Installation

Both the Basic and Advanced installation methods offer the option of creating a starter database during the installation process. If you choose to create a starter database, Oracle Universal Installer uses Database Configuration Assistant (DBCA) to create it.

DBCA is an administrative tool that is provided with Oracle Database. If you do not create a database during the installation process, you can use DBCA to create it afterwards. See "Database Configuration Assistant" on page 2-3.

Preparing for Installation

Several preliminary steps are required before you begin the installation process.

- Verify that your computer meets the minimum hardware and software requirements
- Ensure that you have administrative privileges on your computer
- If you have other Oracle products installed on your computer, stop the Oracle services
- If your computer has a dynamic IP address, install a loopback adapter
- If you are installing from a network drive, use Windows File Manager to map that drive to your computer.

Hardware Requirements

Verify that your computer meets the following requirements:

- Processor: 550 MHz minimum
- Available hard disk space: 2 GB, including 125 MB temp disk space
- RAM: 256 MB minimum, 512 MB recommended
- Virtual memory: double the amount of RAM

Video adapter: 256 color

Software Requirements

The following operating systems are among those supported for Oracle Database:

- Windows XP Professional
- Windows 2000 with service pack 1 or higher

Preexisting Oracle Products

If any Oracle products are already installed on your computer, you must disable them before installing Oracle Database. You need to stop any Oracle services that are currently running, and you need to delete the ORACLE_HOME environment variable if it is defined.

Stop Oracle Services

- From the Control Panel, choose Administrative Tools, then Services.
- On the Services page, scroll down the list of names and locate those that begin with Oracle. Select those services and choose **Stop**.

See "Oracle Services" on page 2-3 for a list of Oracle services.

Delete the ORACLE_HOME Environment Variable

- **1.** From the Control Panel, choose **System**.
- On the Advanced tab of the System Properties page, choose **Environment** Variables.
- 3. If ORACLE_HOME is included in the list of System Variables, select it and click Delete.

Dynamic IP

Dynamic Host Configuration Protocol (DHCP) assigns dynamic IP addresses on a network. Dynamic addressing allows a computer to have a different IP address each time it connects to the network.

If you plan to install Oracle Database on a computer that uses the DHCP protocol, you need to install a loopback adapter to assign a local IP address to your computer.

You can determine if a loopback adapter is already installed on your Windows computer by opening a command window, navigating to the system drive (typically c:), and executing the following command.

ipconfig /all

You can find complete instructions for installing a loopback adapter in the Oracle Database Installation Guide for Microsoft Windows (32-Bit)

Installing the Data Mining Option

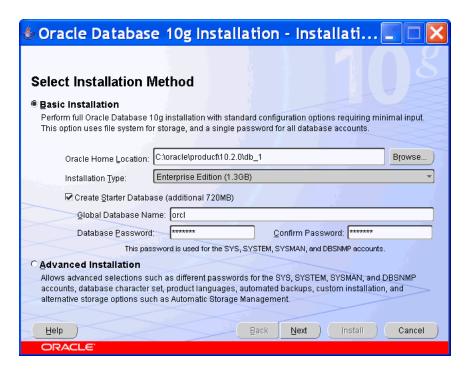
The simplest way to install Oracle Database is to perform a Basic installation. This installation method provides the Data Mining option and the sample schemas by default.

Follow these steps to perform a Basic installation with a starter database on a Windows platform:

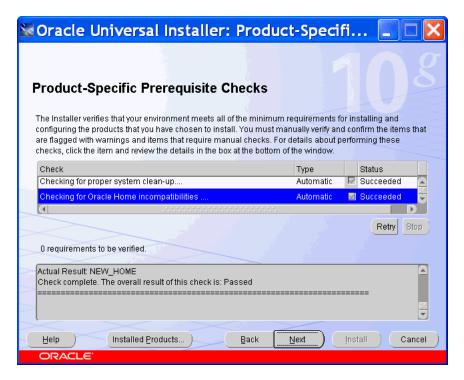
- 1. Check the pre-installation requirements described in "Preparing for Installation" on page 1-3.
- Logon as Administrator.
- From the Database installation directory, run SETUP.EXE.

Oracle Universal Installer opens and displays the Select Installation Method page. Choose Basic Installation. Specify the Oracle home directory. Choose Enterprise **Edition** as the Installation Type. Check the **Create Starter Database** box. Provide a global database name and a password for the system accounts.

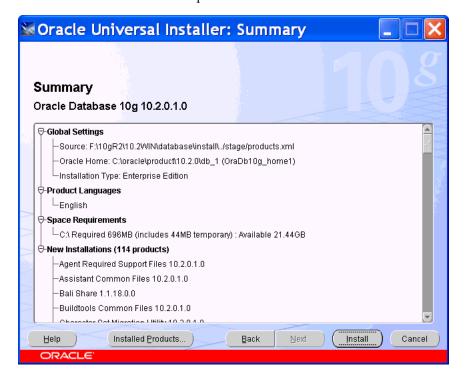
Note: You will have the opportunity to change the passwords for the system accounts at a later time.



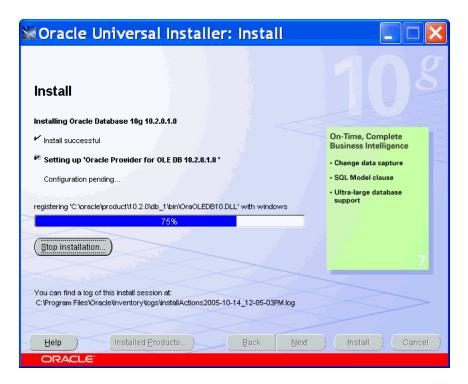
On the Product-Specific Prerequisite Checks page, verify that all checks succeeded. If any checks failed, then you must correct the problem before proceeding.



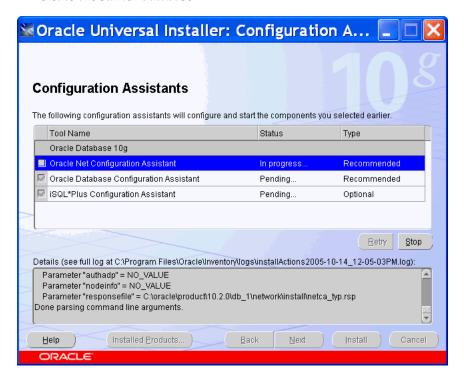
The Summary page displays the settings and components for the installation. Click the **Install** button to proceed.



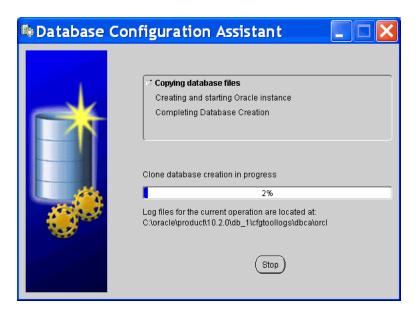
The Installer proceeds with the installation.



After installing the software, the Installer invokes the Configuration Assistants to create the starter database.

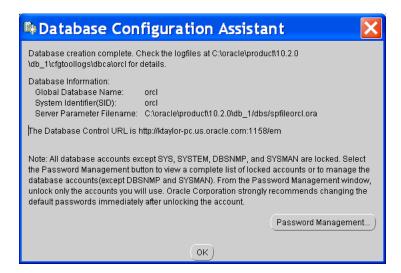


The Database Configuration Assistant copies the files for the starter database.



The Database Configuration Assistant displays information about the starter database. Click the Password Management button. Unlock the DMSYS account and any other accounts that you plan to use. If you plan to use the demo programs, you must unlock the SH account. Click **OK**.

> **Tip:** In the Password Management dialog, you can specify new passwords for the system accounts.



10. The final screen displays the URLs for Enterprise Manager Database Control and for iSQL*Plus. See "Remote Administration" on page 2-4.

Tip: Print a screen shot of this page or write down the information it contains for future reference.

Click Exit to exit the Installer.

Upon completion, the Installer displays the login page for the Database Control.

You can verify your Data Mining installation by querying the V\$OPTION view, as described in "Verifying Installation of Data Mining Components" on page 1-14.

Note: In Oracle Database 10.1, Data Mining was installed in ORACLE HOME\dm. In Oracle 10g Release 2 (10.2), Data Mining is installed in ORACLE HOME\RDBMS.

Adding the Data Mining Option to a Database

Once you have installed the Oracle Database software, you can build databases as needed. You might build a database without the Data Mining option but later decide to add it. To add the Data Mining option, use Database Configuration Assistant (DBCA).

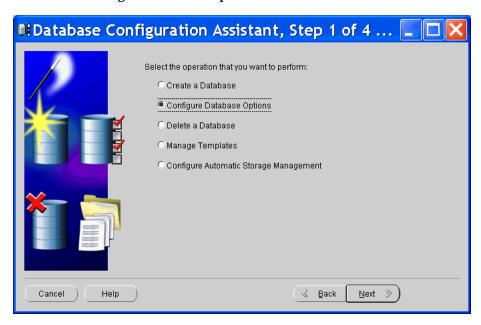
Follow these steps:

- 1. Logon as Administrator.
- Stop any Oracle services that are currently running. See "Stop Oracle Services" on page 1-4.
- **3.** Open DBCA from the Windows Start menu.

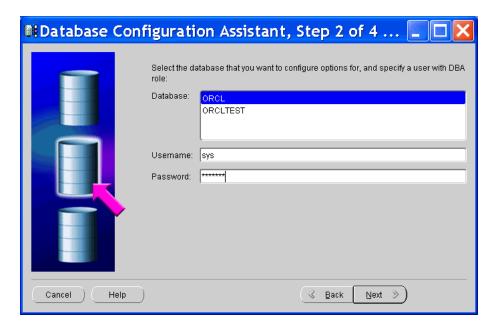
Click Start > Programs > Oracle- oracle_home > **Configuration and Migration Tools > Database Configuration Assistant.**

The Welcome page appears. Click **Next**.

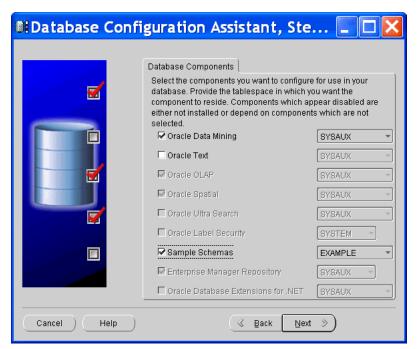
Choose Configure Database Options.



Choose the database to configure, and provide system credentials.



Select **Oracle Data Mining**. Select **Sample Schemas**, if the samples are not already installed and you wish to run the demo programs.



7. Specify dedicated or shared server mode. Click the **Finish** button.

Removing the Data Mining Option

You can use Oracle Universal Installer to remove components of an Oracle installation. If you wish to deinstall the Data Mining option, take these steps:

- Logon as Administrator.
- Stop any Oracle services that are currently running. (See "Stop Oracle Services" on page 1-4.)

3. Open Universal Installer from the Windows Start menu.

Click Start > Programs > Oracle_oracle_home > Oracle Installation Products > Universal Installer.

- **4.** On the Welcome page, click **Deinstall Products**.
- In the Inventory box, choose **Oracle Data Mining RDBMS Files** and click Remove.
- **6.** Oracle Universal Installer removes the Data Mining option and returns to the Inventory page. Click **Close**. To exit the Installer, click **Cancel**.

Installing the Data Mining Scoring Engine

The Data Mining Scoring Engine scores data using models that were created by Oracle Data Mining. You should only install the Data Mining Scoring Engine on computers where you intend to apply models that were built on another computer system. For information on importing models from another database, see "Exporting and Importing Data Mining Models" on page 2-11.

The Scoring Engine is an Enterprise Edition component that can be installed only during a Custom installation.

> In most cases, you will find no advantage in installing the Scoring Engine. The Data Mining option, which is included by default in Oracle Database Enterprise Edition, provides the full range of data mining functionality, including model scoring.

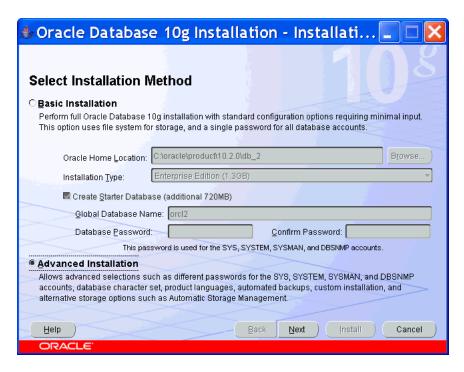
> The Data Mining option and the Scoring Engine are mutually exclusive. When you install one, you automatically deinstall the other. Oracle Data Mining is the only database option affected by installation of the Scoring Engine.

Performing a Custom Installation

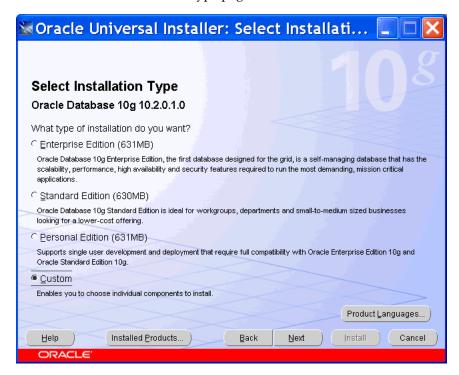
Follow these steps to start the installation of Oracle Database with the Data Mining Scoring Engine:

- Review "Preparing for Installation" on page 1-3.
- Logon as Administrator.
- **3.** From the Database installation directory, run SETUP. EXE.

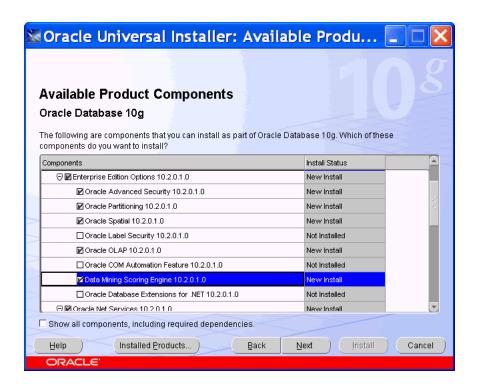
Oracle Universal Installer opens and displays the Select Installation Method page. Choose **Advanced Installation** and click **Next**.



On the Select Installation Type page, choose **Custom**. Click **Next**.



- On the Specify Home Details page, specify the location of a new Oracle home directory.
- **6.** On the Available Product Components page, select **Data Mining Scoring Engine** along with the other components to include in your database. Click Next to continue the installation process.



Adding the Scoring Engine to an Existing Installation

To add the Data Mining Scoring Engine to an existing installation of Oracle Database, take these steps:

- 1. Logon as Administrator.
- Stop any Oracle services that are currently running. (See "Stop Oracle Services" on page 1-4.).
- **3.** Open Oracle Universal Installer from the Windows Start menu.
 - Click Start > Programs > Oracle_oracle_home > Oracle Installation Products > Universal Installer.
- **4.** Choose **Advanced** for the installation method and **Custom** for the installation type.
- 5. On the Specify Home Details page, select the Oracle home directory in which you previously installed Oracle Database 10g Release 2 (10.2). Do not rely on the default setting to be correct.
- **6.** On the Available Product Components page, select **Data Mining Scoring Engine**. Click **Next** to continue the installation process.

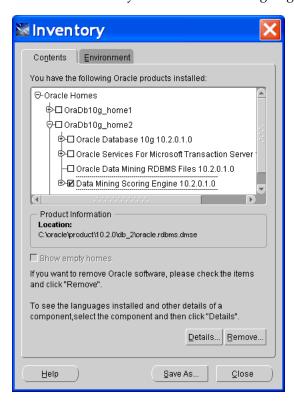
Removing the Scoring Engine

If you wish to deinstall the Scoring Engine, take these steps.

- **1.** Logon as Administrator.
- 2. Stop any Oracle services that are currently running. (See "Stop Oracle Services" on page 1-4.)
- **3.** Open Universal Installer from the Windows Start menu.

Click Start > Programs > Oracle—oracle_home > Oracle Installation Products > Universal Installer.

- 4. On the Welcome page, click **Deinstall Products**.
- **5.** In the Inventory box, choose **Data Mining Scoring Engine** and click **Remove**. Note that **Oracle Data Mining RDBMS Files** is a separate component. It is not removed when you remove the Scoring Engine.



Oracle Universal Installer removes the Scoring Engine and returns to the Inventory page. Click **Close**. To exit the Installer, click **Cancel**.

Verifying Installation of Data Mining Components

To verify the status of Data Mining in your database, start SQL*Plus and log in as SYSDBA.

If Data Mining is installed, then the following SQL query returns TRUE.

```
SELECT VALUE FROM V$OPTION WHERE PARAMETER = 'Data Mining';
```

If the Data Mining Scoring Engine is installed, then the next SQL query returns TRUE.

```
SELECT VALUE FROM V$OPTION WHERE PARAMETER = 'Data Mining Scoring Engine';
```

For more detailed information, query DBA REGISTRY with a command such as this one.

SQL> SELECT COMP ID, VERSION, STATUS FROM DBA REGISTRY WHERE COMP ID = 'ODM';

COMP_ID	VERSION	STATUS
ODM	10.2.0.0.0	VALID

Tips for Linux Installations

The step-by-step instructions in this guide are for installations on a Windows platform. The following are some tips to help you with the differences between the Linux and Windows platforms.

- Follow the instructions in the installation guide for your platform. Be sure to set system variables and perform any other pre-installation tasks.
- During the installation, you must run some SQL scripts as root, so make sure that you have root access.
- To run Oracle Universal Installer, type runInstaller from the installation disk or directory.
- Create a script for setting the appropriate environment variables, or include the settings in the initialization script for your operating system login (such as .profile or .cshrc). Perform any other post-installation tasks.

The following is a sample script for setting environment variables:

```
setenv ORACLE ROOT /dat1/10gR2
setenv ORACLE BASE ${ORACLE ROOT}
setenv ORACLE HOME ${ORACLE_BASE}/oracle/product/10.2.0/db_1
setenv ORACLE PORT 1521
setenv ORACLE SID rel10g
setenv TNS ADMIN ${ORACLE HOME}/network/admin
setenv PATH ${ORACLE HOME}/bin:${ORACLE HOME}:$PATH
```

See Also:

- Oracle Database Quick Installation Guide for Linux x86 for a basic installation
- *Oracle Database Installation Guide for Linux x86* for a custom installation, upgrade, or other variation

Upgrading from Oracle Database 10g Release 1 (10.1)

If Oracle Database 10g Release 1 is installed on your computer, then install Oracle Database 10g Release 2 into a separate Oracle home. A new version of the Data Mining option is installed along with the other components of Oracle Database.

During installation of release 10.2, Oracle Universal Installer will detect 10.1 databases and will prompt you to upgrade them. For upgrading, the Installer will open the Database Upgrade Assistant (DBUA). It automatically upgrades these Data Mining components:

- **Databases**
- DMSYS repository schema
- PL/SQL data models in user schemas

See Also: Oracle Database Upgrade Guide for complete upgrade instructions

Post-Upgrade Steps for Data Mining

The data mining objects in these schemas require additional upgrade steps after upgrading the database:

- DMSYS Repository Schema. The odmu101s.sql script removes redundant release 10.1 objects.
- User Schemas. The odmu101a.sql script re-creates the views of the Sales History sample schema for users who wish to run the Data Mining demo programs described in Chapter 4.

To upgrade the DMSYS repository, follow these steps:

- 1. Open SQL*Plus and connect to the database as the SYS user or another privileged account.
- **2.** Verify that the compatibility mode is set to 10.2.*x*.*x* (such as 10.2.0.0).

```
SELECT name, value, description FROM v$parameter
        WHERE name='compatible';
```

If the compatibility model is incorrect, then edit the init.ora file for the database.

3. Run the odmu101s.sql script.

Issue this command on Windows platforms:

```
%ORACLE HOME%\rdbms\admin\odmu101s.sql
```

Or this command on Linux platforms:

```
$ORACLE_HOME/rdbms/admin/odmu101s.sql
```

To upgrade the Sales History demo views in the schemas of Data Mining users, take these steps for each schema:

- Backup the schema using the expdp command in Oracle Data Pump.
- 2. Open SQL*Plus and connect to the database as the SYS user or another privileged account.
- **3.** Run the odmu101a.sql script, passing it two parameters: *user_name* and password.

Issue this command on Windows platforms:

```
%ORACLE_HOME%\rdbms\admin\odmu101a.sql user_name password
```

Or this command on Linux platforms:

```
$ORACLE HOME/rdbms/admin/odmul0la.sql user name password
```

Upgrading PL/SQL Models

All data mining PL/SQL models residing in user schemas in a release 10.1 database can be upgraded as part of the database upgrade. The upgrade of models is seamlessly integrated into the database upgrade process.

You can also export models from a release 10.1 database and import them into a release 10.2 database. The import process upgrades the models to release 10.2. Use the expdp and impdp command utilities in Oracle Data Pump. The Data Mining option or the Data Mining Scoring Engine option must be installed in the target database.

Note: Java models cannot be upgraded. They must be re-created in Oracle Database 10.2.

The Database Upgrade Assistant invokes procedures in the Data Mining DMP SYS public package for upgrading and downgrading PL/SQL models between release 10.1 and 10.2.

The UPGRADE_MODELS procedure upgrades the Data Mining models from release 10.1 to 10.2, and has this syntax:

```
UPGRADE MODELS (
    to_version
                  IN
                         VARCHAR2);
```

The DOWNGRADE_MODELS procedure downgrades models from release 10.2 to 10.1, and has this syntax:

```
DOWNGRADE MODELS (
    from_version IN
                          VARCHAR2);
```

These procedures are integrated with the Data Mining upgrade and downgrade processes and run as SYS. You do not need to run them manually.

Downgrading Oracle Data Mining

You can downgrade a database from release 10.2 to release 10.1 if the compatibility setting is set to 10.1.x.x. Otherwise, the database will include incompatibilities that will prevent it from being downgraded. The Data Mining option and user-owned PL/SQL models will be downgraded as part of database downgrade.

See Also: Oracle Database Upgrade Guide for complete downgrade instructions

Upgrading from Oracle Database 10g Release 1 (10.1
--

Administering Oracle Database for Data Mining

Because Oracle Data Mining is completely integrated with Oracle Database, you will use Oracle Database tools to administering Data Mining. You can administer Oracle Database locally or from a remote computer with network access.

In this chapter, you will learn about post-installation administrative tasks, such as creating users and exporting and importing models.

This chapter contains the following topics:

- Local Administration on Microsoft Windows
- Local Administration on Linux
- Remote Administration
- Creating Data Mining Users
- **Exporting and Importing Data Mining Models**

Local Administration on Microsoft Windows

Several tools for administrators and application developers are installed along with Oracle Database. For Microsoft Windows platforms, the Start menu contains an Oracle home program group with links to the tools.

Following are descriptions of a few of the basic administrative tools.

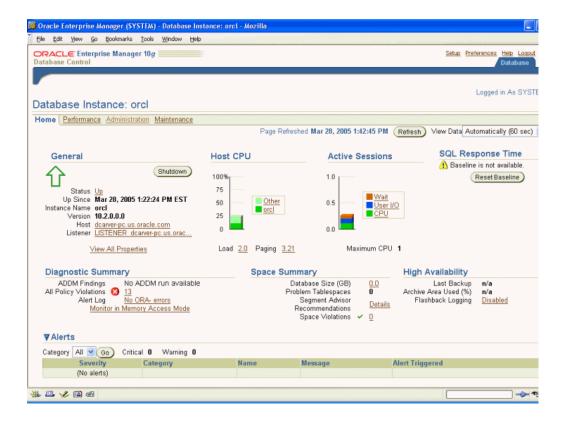
Oracle Enterprise Manager Database Control

Database Control provides a Web-based graphical interface for managing all aspects of Oracle Database.

To open Database Control from the Windows Start menu, click **Start > Programs >** Oracle- *oracle_home* > Database Control- *database_name*.

You can also open Database Control from the URL provided during installation.

The following figure shows the Database Control home page.



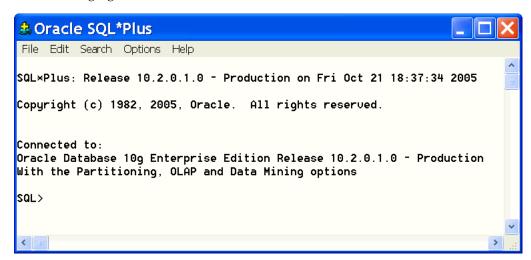
SQL*Plus

SQL*Plus is a command-line interface for the SQL language. You can perform all Oracle administrative tasks using SQL.

To open SQL*Plus from the Windows Start menu, click **Start** > **Programs** > Oracle- oracle_home > Application Development > SQL Plus.

You will be prompted for your user name and password. You must supply a host string only when connecting to a remote computer. The host string takes the form host_name:port:SID, such as myhost:1521:orcl.

The following figure shows the SQL Plus window.



Database Configuration Assistant

Database Configuration Assistant provides a graphical user interface for creating, configuring, and deleting database instances. A single installation of Oracle Database can support numerous individual database instances. You can use Database Configuration Assistant to install the sample schemas if you did not install them with the database.

To open Database Configuration Assistant from the Windows Start menu, click Start > Programs > Oracle- oracle_home > Configuration and Migration Tools > **Database Configuration Assistant.**

Oracle Universal Installer

You can use Oracle Universal Installer to list the Oracle products on your computer or to deinstall them.

To open Oracle Universal Installer from the Windows Start menu, click **Start** > Programs > Oracle_home > Oracle Installation Products > Universal Installer.

You must shut down all databases and supporting services before deinstalling Oracle Database. Refer to the installation guide for your platform for more information.

Oracle Services

The Oracle Database installation creates several services. The following table describes some of them.

Service Name	Description	Usage
OracleService <i>SID</i>	Oracle Database	Enables you to start and stop Oracle Database from the Service window.
OracleHome_NameTNSListener	Oracle Database listener	Enables you to open a connection with Oracle Database from a remote computer.
OracleHome_NameiSQL*Plus	iSQL*Plus application server	Enables you to open iSQL*Plus from a browser.
OracleDBConsoleSID	Oracle Enterprise Manager Database Control console	Enables you to open Database Control from a browser.

To manage them, open Administrative Tools in the Windows Control Panel and choose Services.

Local Administration on Linux

The same tools that are installed locally on a Windows platform are also installed on Linux. You can run the local administrative tools from the shell command line. They are located in \$ORACLE HOME/bin. These are a few of the tools:

- To open SQL*Plus, type sqlplus.
- To open Database Configuration Assistant, type dbca.
- To open Enterprise Manager Database Control, open a browser and type the URL provided during installation.

- To open Oracle Universal Installer, type \$ORACLE HOME/oui/bin/runInstaller.
- To start and stop the various Oracle processes, use these commands:
 - 1snrct1: Oracle Database listener
 - isqlplusctl: iSQL*Plus application server
 - emctl: Oracle Enterprise Manager Database Control console

For descriptions of these tools, refer to "Local Administration on Microsoft Windows" on page 2-1.

Remote Administration

You can open these tools in any browser by typing the URLs listed during installation on the End of Installation page:

- *i*SQL*Plus is a version of SQL*Plus that runs in a browser.
- Enterprise Manager Database Control is the same thin-client application that you access locally.

The administration tools installed with Oracle Database are also installed with Oracle Client. If you are administering a remote database, you must install Oracle Client to obtain the full suite of administration tools. See "Installing Oracle Client" on page 3-4 for information on Oracle Client installation.

Creating Data Mining Users

Anyone who wants to use Oracle Database must have a user name and password. Data Mining users must have several database permissions, plus SELECT access to the tables containing data for analysis. Data mining activities typically require resources for working with large amounts of data.

The examples in this chapter show how to use Database Control or SQL commands to create data mining users. You can cut and paste the SQL commands into SQL*Plus.

Note: In Oracle Database 10.1, a data mining user account, called DMUSER, was provided with the software. In Oracle Database 10g Release 2 (10.2), DMUSER is no longer provided; you must create your own data mining user accounts.

See Also: Chapter 4 if you wish to create a demo user with permissions for running the demo programs.

Creating Tablespaces

All users require a permanent tablespace and a temporary tablespace in which to do their work. Performance may start to degrade if multiple users are sharing the same tablespace while mining large data sets. You can improve performance by creating individual tablespaces for each user.

You must log in with system privileges to create tablespaces.

Using SQL Commands

The following SQL command creates a new permanent tablespace.

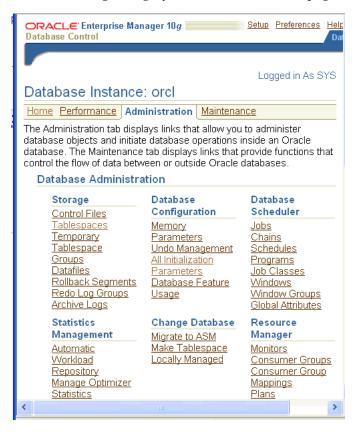
```
CREATE TABLESPACE "ODMPERM" DATAFILE
    'C:\ORACLE\PRODUCT\10.2.0\ORADATA\ORCL\odm1.dbf'
    SIZE 20M REUSE AUTOEXTEND ON NEXT 20M;
```

The next SQL command creates a new temporary tablespace.

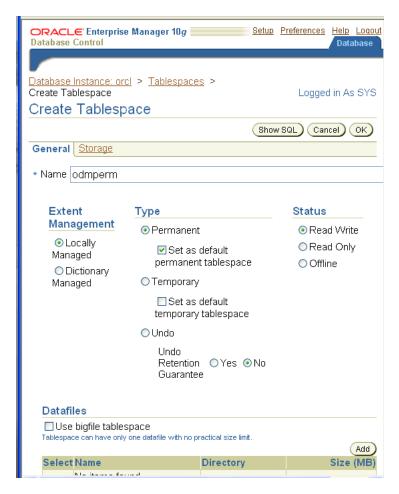
```
CREATE TEMPORARY TABLESPACE "ODMTEMP" TEMPFILE
    'C:\ORACLE\PRODUCT\10.2.0\ORADATA\ORCL\odmtemp.tmp'
    SIZE 20M REUSE AUTOEXTEND ON NEXT 20M;
```

Using Enterprise Manager Database Control

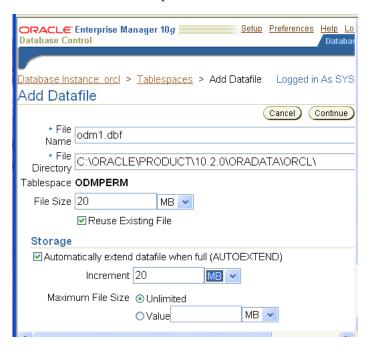
- From the Database Control home page, select the Administration tab.
- In the Storage category on the Administration page, choose **Tablespaces**.



- On the Tablespaces page, click the **Create** button.
- On the General tab of the Create Tablespace page, provide a name for the tablespace. Specify whether it is temporary or permanent, and set it as the default. To create the datafile, click the **Add** button.



On the Add Datafile page, provide a file name for the tablespace. Specify the file size and click the **Reuse Existing File** box. In the Storage section, specify **AUTOEXTEND** and provide the increment size. Click **Continue**.



6. On the Create Tablespace page, click OK.

Creating Users

Data mining users that will mine large data sets should have personal tablespaces, as described in "Creating Tablespaces" on page 2-4.

Access Rights

Data mining users require several CREATE privileges. For text mining, users must also have access to the Oracle Text package ctxsys.ctx ddl. The following privileges are required.

```
CREATE PROCEDURE
CREATE SESSION
CREATE TABLE
CREATE SEQUENCE
CREATE VIEW
CREATE JOB
CREATE TYPE
CREATE SYNONYM
EXECUTE ON ctxsys.ctx ddl
```

Data mining users must also have SELECT privileges on the data being mined.

See Also: Chapter 4 for information about the shgrants script, which grants privileges to a demo user.

Users who want to export and import Data Mining models need additional access rights, as described in "Exporting and Importing Data Mining Models" on page 2-11.

You must log in with system privileges to create users.

Using SQL

The following command creates a user named DMUSER1 with the password change_now, and provides default access to two personal tablespaces.

```
CREATE USER dmuser1 IDENTIFIED BY change_now
    DEFAULT TABLESPACE odmperm
    TEMPORARY TABLESPACE odmtemp
    QUOTA UNLIMITED on odmperm;
```

The following commands in a stored procedure grant the required privileges to the user DMUSER1.

```
GRANT create procedure to DMUSER1
GRANT create session to DMUSER1
GRANT create table to DMUSER1
GRANT create sequence to DMUSER1
GRANT create view to DMUSER1
GRANT create job to DMUSER1
GRANT create type to DMUSER1
```

```
GRANT create synonym to DMUSER1
GRANT execute on ctxsys.ctx_ddl to DMUSER1
```

Unless a user owns the data being analyzed, you must grant access rights to that data using a SQL command like this one.

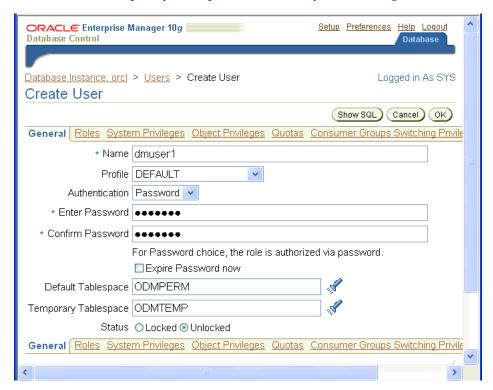
```
GRANT SELECT ON owner.tablename TO user
```

For example, the following SQL command grants SELECT access to the EMPLOYEES table in the sample HR schema to DMUSER1.

```
GRANT SELECT ON hr.employees TO DMUSER1;
```

Using Enterprise Manager Database Control

- On the Database Control home page, select the Administration tab.
- In the Users & Privileges category on the Administration page, choose **Users**.
- On the Users page, click **Create**.
- **4.** On the Create User page, provide the user name and password, and specify default and temporary tablespaces. Select the System Privileges tab.



- On the System Privileges page, choose **Edit List**.
- On the Modify System Privileges page, use the arrows to move the required privileges from the Available box to the Selected box. Click **OK**.



On the Create User page, select the Object Privileges tab to grant access to the mining data. In the Select Object Type drop down list, choose Table. Click Add.



8. On the Add Table Object Privileges page, choose the table where the mining data is stored and grant SELECT access to it. The following example provides the user with SELECT access to the sample HR. COUNTRIES table. Click **OK**.



- Repeat steps 7 and 8, choosing Package instead of Table. Grant EXECUTE access to the CTXSYS.CTX_DLL package. Click **OK**.
- 10. Click Apply.
- 11. You can view the definition of the user you have created by clicking the View button.



About the DMSYS Schema

Information about all models created in a database is stored in tables owned by the DMSYS user. During a typical installation, the DMSYS user has SYSAUX defined as its default tablespace.

Important: Do not delete, truncate, or modify the tables in the DMSYS schema. They support the data mining activities of all users in the database.

Exporting and Importing Data Mining Models

You can export data mining models to flat files to back up work in progress or to move models to a different instance of Oracle Database Enterprise Edition (such as from a development database to a production database). All methods for exporting and importing models are based in Oracle Data Pump technology.

Oracle Data Pump consists of two command-line clients and two PL/SQL APIs. The command-line clients, expdp and impdp, provide an easy-to-use interface to the Data Pump export and import utilities. The Data Mining APIs also use the Data Pump export and import utilities.

You can export and import models at different levels, depending on your access rights in the database:

- Database. When a DBA exports a full database using expdp, all data mining models in the database are exported. The impdp utility imports all the models with the other objects in the database.
- Schema. When a DBA or an individual user exports a schema using expdp, all the data mining models in the schema are exported. Likewise, impdp imports all the models with the other objects in the schema.
- Models Only. The Data Mining APIs contain utilities for exporting and importing either all Data Mining models in a schema or models that match specific criteria.

The Data Pump export utility writes the tables and metadata that constitute a model to a dump file set, which consists of one or more files. The Data Pump import utility retrieves the tables and metadata from the dump file and restores them to the target database. Because the expdp and impdp clients and the Data Mining APIs use the Data Pump export and import utilities, you can use the APIs to extract individual models from a dump file of a schema or database.

Note that the older exp and imp database utilities do not export or import data mining models.

See Also:

- Oracle Database Utilities for a complete discussion of Oracle Data Pump and the expdp and impdp utilities
- Oracle Database PL/SQL Packages and Types Reference for detailed information about the export and import procedures in the DBMS DATA MINING package.
- Oracle Data Mining Java API Reference for information about the export and import classes in the Oracle Data Mining Java API.

Prerequisites

To export and import Data Mining models, you must have read and write access to a directory object, and you may need additional database permissions.

Directory Objects

A directory object is a logical name in the database for a physical directory on the host computer. Without read and write access to a directory object, you cannot access the host computer file system from within Oracle Database.

You must have the CREATE ANY DIRECTORY privilege to create directory objects.

The following SQL command creates, or re-creates if it already exists, a directory object named DMTEST. The file system directory (in this example, C:\ORACLE\PRODUCT\10.2.0\DMINING) must already exist and have shared read/write access rights granted by the operating system.

CREATE OR REPLACE DIRECTORY dmtest AS 'c:\oracle\product\10.2.0\dmining';

This SQL command gives user DMUSER1 both read and write access to DMTEST.

GRANT ALL ON DIRECTORY dmtest TO dmuser1;

For more information about creating database directories, refer to the CREATE DIRECTORY and GRANT commands in the *Oracle Database SQL Reference*.

Additional Database Privileges

You may need special privileges in the database to take full advantage of all Data Pump features, such as importing models and other objects into a different schema. These privileges are granted by the EXP FULL DATABASE and IMP FULL DATABASE roles.

You do not need these roles to export models from your own schema. To import models, you must have the same database roles or be as privileged as the user who created the dump file set. Otherwise, you need the IMP FULL DATABASE role.

Privileged users (such as SYS or a user with the DBA role) have sufficient access rights and do not need these additional roles.

The following SQL commands grant these roles to DMUSER1:

```
GRANT EXP FULL DATABASE TO dmuser1;
GRANT IMP FULL DATABASE TO dmuser1;
```

PL/SQL APIs for Exporting and Importing Models

The DBMS DATA MINING PL/SQL package contains these two procedures:

- EXPORT MODEL
- IMPORT MODEL

For more information about these procedures, refer to the Oracle Database PL/SQL Packages and Types Reference.

Java APIs for Exporting and Importing Models

Oracle Database implements the industry-standard Java Data Mining (JDM) API Specification, which includes these two interfaces:

- javax.datamining.task.ExportTask
- javax.datamining.task.ImportTask

For more information about the standard JDM API, refer to the Java Help for the JSR-73 Specification, which is available on the Oracle Technology Network at

http://www.oracle.com/technology/products/bi/odm/JSR-73/index.html

Tables Created By Exporting and Importing Models

Two tables are created in the user's schema by the Data Mining export and import utilities:

- DM\$P MODEL EXPIMP TEMP. Used for internal purposes during export and import, and provides a job history.
- DM\$P MODEL TABKEY TEMP. Used only for internal purposes during export and import.

Do not alter these tables. However, you may drop them when no export or import job is running. The utilities will re-create them for the next job.

Example: Exporting and Importing Models

This example creates a dump file with three models and imports the models from the dump file.

Exporting Models from the DMUSER Schema

The following command exports all models from DMUSER, who is currently connected to the database in SQL*Plus.

```
SQL> EXECUTE DBMS_DATA_MINING.EXPORT_MODEL('allmodels.dmp','DMTEST');
PL/SQL procedure successfully completed.
```

An export or import creates a log file in the same directory as the dump file. Error messages are returned to the current output device (such as the screen), and the log file may provide additional information.

This command was successful and creates two files in the DMTEST directory:

- A dump file named allmodels01.dmp (note the 2-digit suffix added to the name)
- A log file with a default name of DMUSER exp 4589.log

For detailed information about the default names of files, see the DBMS DATA MINING package in the Oracle Database PL/SQL Packages and Types Reference.

You can view the log file using a system command or editor. You must know the path of the physical directory in order to locate the file.

DMUSER exp 4589.log lists the three Data Mining models that were in the schema, plus additional objects as shown here:

```
Starting "DMUSER"."DMUSER_exp_45": DM_EXPIMP_JOB_ID=45
Estimate in progress using BLOCKS method...
Processing object type TABLE_EXPORT/TABLE/TABLE_DATA
Total estimation using BLOCKS method: 1.062 MB
>>> . . exported Data Mining Model "DMUSER". "ABN CLAS SAMPLE"
>>> . . exported Data Mining Model "DMUSER". "ASSOCIATION RULES SAMPLE"
>>> . . exported Data Mining Model "DMUSER". "NAIVE BAYES SAMPLE"
Processing object type TABLE EXPORT/TABLE/PROCACT INSTANCE
Processing object type TABLE_EXPORT/TABLE/TABLE
Processing object type TABLE_EXPORT/TABLE/GRANT/OWNER_GRANT/OBJECT_GRANT
Processing object type TABLE EXPORT/TABLE/INDEX/INDEX
Processing object type TABLE EXPORT/TABLE/CONSTRAINT/CONSTRAINT
Processing object type TABLE EXPORT/TABLE/INDEX/STATISTICS/INDEX STATISTICS
Processing object type TABLE EXPORT/TABLE/INDEX/FUNCTIONAL AND BITMAP/INDEX
Processing object type
TABLE EXPORT/TABLE/INDEX/STATISTICS/FUNCTIONAL AND BITMAP/INDEX STATISTICS
Processing object type TABLE EXPORT/TABLE/STATISTICS/TABLE STATISTICS
. . exported "DMUSER"."DM$POASSOCIATION_RULES_SAMPLE" 7.640 KB 15 rows
. . exported "DMUSER"."DM$PONAIVE_BAYES_SAMPLE" 18.35 KB
                                                                            219 rows
. . exported "DMUSER"."DM$P1ABN_CLAS_SAMPLE" 6.945 KB 2 rows . . exported "DMUSER"."DM$P1NAIVE_BAYES_SAMPLE" 5.929 KB 2 rows
. . exported "DMUSER"."DM$P2ASSOCIATION RULES SAMPLE" 6.210 KB
                                                                           11 rows
. . exported "DMUSER"."DM$P3ASSOCIATION RULES SAMPLE" 6.179 KB 18 rows
. . exported "DMUSER"."DM$P4ASSOCIATION RULES SAMPLE" 5.492 KB
                                                                           26 rows
. . exported "DMUSER"."DM$P5ABN_CLAS_SAMPLE"
                                                           5.304 KB
                                                                            2 rows
. . exported "DMUSER"."DM$P5NAIVE_BAYES_SAMPLE"
. . exported "DMUSER"."DM$P6ABN_CLAS_SAMPLE"
                                                            5.984 KB
                                                                           27 rows
                                                            16.47 KB
                                                                            34 rows
. . exported "DMUSER"."DM$P7ABN_CLAS_SAMPLE"
. . exported "DMUSER"."DM$P8ABN_CLAS_SAMPLE"
                                                            7.007 KB
                                                                              5 rows
                                                           5.414 KB
                                                                              5 rows
. . exported "DMUSER"."DM$P8ASSOCIATION_RULES_SAMPLE" 5.335 KB
                                                                              3 rows
. . exported "DMUSER"."DM$P8NAIVE_BAYES_SAMPLE" 5.359 KB
. . exported "DMUSER"."DM$PEABN_CLAS_SAMPLE" 9.093 KB
. . exported "DMUSER"."DM$PENAIVE_BAYES_SAMPLE" 8.742 KB
. . exported "DMUSER"."DM$P_MODEL_EXPIMP_TEMP" 6.273 KB
                                                                             3 rows
                                                            9.093 KB 116 rows
                                                           8.742 KB 116 rows
                                                           6.273 KB 10 rows
```

```
. . exported "DMUSER"."DM$PEASSOCIATION RULES SAMPLE"
                                                  0 KB
                                                             0 rows
Master table "DMUSER"."DMUSER_exp_45" successfully loaded/unloaded
************************
Dump file set for DMUSER.DMUSER exp 45 is:
 /dat2/10gR2/oracle/product/10.2.0/db 1/dmtest/allmodels01.dmp
Job "DMUSER". "DMUSER exp 45" successfully completed at 08:40:08
```

Importing Models Into the Same Schema

DMUSER can restore these models from the dump file at a later date if, for whatever reason, he or she wants to revert to this version of the models. Note that an import will not overwrite an existing model with the same name unless the model is incomplete or corrupted.

The following command restores all models from the dump file to the DMUSER schema:

```
SQL> EXECUTE DBMS DATA MINING.IMPORT MODEL('allmodels01.dmp','DMTEST');
```

Importing Models Into a Different Schema

A user with the necessary privileges can load the models from a dump file into a different schema. In the next example, the SYSTEM user issues the following command, which loads the three models into the SCOTT schema:

```
SQL> EXECUTE DBMS DATA MINING.IMPORT MODEL('allmodels01.dmp', 'DMTEST', null,
null, null, 'toscott', 'DMUSER:SCOTT');
```

This import command specifies to scott.log as the name of the log file; the .log extension is added automatically to the name. The log file shows the names of the imported models and supporting metadata.

```
Master table "SYSTEM"."toscott" successfully loaded/unloaded
Starting "SYSTEM"."toscott": DM EXPIMP JOB ID=51 DM SELECT IMPORT
Processing object type TABLE EXPORT/TABLE/PROCACT INSTANCE
>>> . . imported Data Mining Model "SCOTT". "ABN CLAS SAMPLE"
>>> . . imported Data Mining Model "SCOTT". "ASSOCIATION RULES SAMPLE"
>>> . . imported Data Mining Model "SCOTT". "NAIVE BAYES SAMPLE"
Processing object type TABLE EXPORT/TABLE/TABLE
Processing object type TABLE EXPORT/TABLE/TABLE DATA
Processing object type TABLE EXPORT/TABLE/GRANT/OWNER GRANT/OBJECT GRANT
Processing object type TABLE EXPORT/TABLE/INDEX/INDEX
Processing object type TABLE EXPORT/TABLE/CONSTRAINT/CONSTRAINT
Processing object type TABLE EXPORT/TABLE/INDEX/STATISTICS/INDEX STATISTICS
Processing object type TABLE EXPORT/TABLE/STATISTICS/TABLE STATISTICS
Job "SYSTEM". "toscott" completed with 1 error(s) at 09:08:12
```

Exi	oortina	and	oaml	rtina	Data	Mining	Models

Installing Tools for Data Mining

Oracle provides two tools to assist analysts in data mining activities: Oracle Data Miner and Oracle Spreadsheet Add-In for Predictive Analytics.

Users must have the database and object permissions described in Chapter 2 to use these tools.

> **Note:** Oracle Data Mining also supports PL/SQL and Java APIs. Application developers can use these APIs to create their own data mining tools. The data mining APIs are documented in Oracle Data Mining Application Developer's Guide.

This chapter contains the following topics:

- Oracle Data Miner
- Oracle Spreadsheet Add-In for Predictive Analytics

Oracle Data Miner

Oracle Data Miner is a graphical user interface for Oracle Data Mining. Oracle Data Miner's easy-to-use wizards guide you through the data preparation, data mining, model evaluation, and model scoring process.

Oracle Data Miner is supported on Windows 2000, Windows XP Professional Edition, and Linux.

Installing Oracle Data Miner on Windows

To install Oracle Data Miner on a Microsoft Windows platform, take these steps:

- Download Oracle Data Miner 10.2 from the Oracle Web site at http://www.oracle.com/technology/products/bi/odm/index.html.
- 2. Open the ZIP file and extract all files to an empty directory (such as C:\ODMINER). Be sure to use folder names so that the files retain their original organization in subfolders.
- Create a Windows shortcut to BIN\ODMINERW. EXE, and drag the shortcut to your Windows desktop for easy access.

Installing Oracle Data Miner on Linux

Follow these instructions for installing Oracle Data Miner on Linux.

Prerequisite

Oracle Data Miner requires Java 1.4.2, which is installed with Oracle Database and Oracle Client. If you are installing Oracle Data Miner on a computer where neither of these products is installed, then check the Java version with this command at the operating system prompt:

```
java -version
```

If Java 1.4.2 is not already installed, you can download it from http://www.java.com.

Installation Steps

To install Oracle Data Miner on a Linux or Unix platform, take these steps:

- Download Oracle Data Miner from the Oracle Web site at http://www.oracle.com/technology/products/bi/odm/index.html.
- Open the ZIP file and extract all files to an empty directory. The following command creates a directory named odminer and inflates the files into it.

```
unzip odminer.zip -d odminer
```

3. Grant execution permission to bin/odminer with a command such as this:

```
chmod +x odminer/bin/odminer
```

4. Run the odminer executable from the bin subdirectory.

```
chmod odminer/bin
odminer
```

Connecting Oracle Data Miner to Oracle Database

To create a connection from Oracle Data Miner to Oracle Database, take these steps:

- 1. Open Oracle Data Miner.
- **2.** On the Oracle Data Miner dialog box, click **New**.
- On the New Connection dialog box, enter the information for connecting to Oracle Database.



The new connection will appear in the Oracle Data Miner dialog box, where you can use it to connect to Oracle Database.



The following figure shows the main page of Oracle Data Miner after opening a new connection.

For documentation to assist you in using Oracle Data Miner, see the online help. Refer also to *Oracle Data Mining Concepts*.

Oracle Spreadsheet Add-In for Predictive Analytics

The Oracle Spreadsheet Add-In for Predictive Analytics adds predictive analytics features to Microsoft Excel. Using simple "one click" Predict and Explain predictive analytics operations, Excel users can mine data stored in Excel or in Oracle Database.

Predictive analytics provide automated methodologies that simplify data mining. Using Predictive Analytics, many more users can harness the power of Oracle Data Mining without possessing the knowledge of advanced data analysts.

Installation Steps

The Spreadsheet Add-In requires several components on your PC:

- Microsoft Excel 2000, 2002, or 2003
- Oracle Objects for OLE

The Spreadsheet Add-In also requires access to an instance of Oracle Database installed with the Data Mining option.

Before you can use the Spreadsheet Add-In, you must take several steps to prepare your PC:

- 1. Install the Add-In in Excel. See "Adding the Spreadsheet Add-In to Excel" on page 3-4.
- 2. Install Oracle Client. This process includes the installation of Oracle Objects for OLE and Oracle Net Configuration Assistant, both of which are required by the Add-In. See "Installing Oracle Client" on page 3-4.
- **3.** Create an Oracle Net Service Name to manage the connection between the client and Oracle Database. See "Creating an Oracle Net Service Name" on page 3-6.

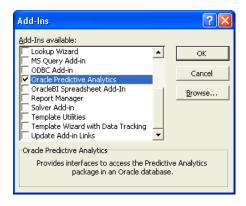
Adding the Spreadsheet Add-In to Excel

To install the Spreadsheet Add-In on your personal computer, take these steps:

- Download the Spreadsheet Add-In for Predictive Analysis from the Oracle Web
 - http://www.oracle.com/technology/products/bi/odm/index.html.
- 2. Open the ZIP file and extract the file named Predictive Analytics.xla to the Microsoft Office Library directory. The library has a path such as this one:
 - C:\Program Files\Microsoft Office\Dibrary

Note: If the scripts dbmsdmpa.sql, prvtdmpa.plb, and Predictive Analytics Demo.sql are included in the ZIP file, you can disregard them. These scripts are used with Oracle Database 10.1; they cannot be used with Oracle Database 10g Release 2 (10.2).

- Open Excel and click **Tools** > **Add-Ins**.
- Select Oracle Predictive Analytics from the Add-Ins dialog box, as shown in the following figure.



The OraclePA menu is added to the Excel toolbar.

Installing Oracle Client

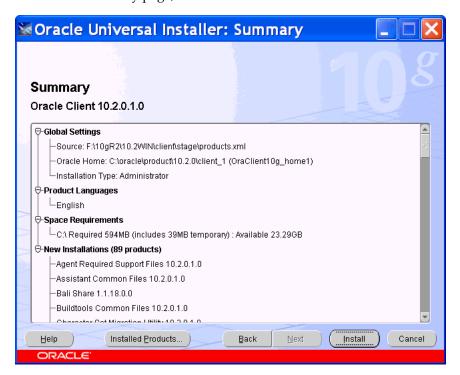
Oracle Spreadsheet Add-In for Predictive Analytics requires the presence of Oracle Objects for OLE. To obtain this component, you must install Oracle Client on the computer where Excel and the Spreadsheet Add-In are installed.

Use the following steps to install Oracle Client on a Windows platform:

- **1.** Logon as Administrator.
- From the Client installation directory, run SETUP.EXE. Oracle Universal Installer opens and displays the Welcome page. Click Next.
- **3.** On the Select Installation Type page, choose **Administrator** and click **Next**.

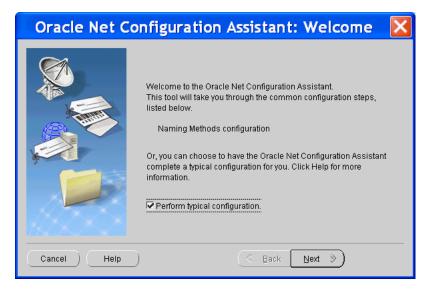


- On the Specify Home Details page, provide the path of a home directory for the Oracle Client installation.
- 5. On the Product-Specific Prerequisite page, verify that all checks succeeded. If any checks failed, then you must correct the problem before proceeding.
- **6.** On the Summary page, click the **Install** button.



The Installer displays the progress of the installation. When you click **Next**, the Configuration Assistants page is displayed.

8. Oracle Net Configuration Assistant starts and displays the Welcome page. Choose Perform Typical Configuration.



Oracle Net Configuration Assistant creates a simple connection using the Easy Connect naming method. This method enables clients to connect to a remote database server without any configuration. Clients specify a SQL CONNECT statement using a simple TCP/IP address, identified by a host name and an optional port number and service name.

CONNECT username/password@host[:port] [/servicename]

See Also: "Creating an Oracle Net Service Name" and *Oracle* Database Net Services Administrator's Guide

9. When the Oracle Net Configuration process is complete, click **Finish**.

Creating an Oracle Net Service Name

The Spreadsheet Add-In requires an Oracle Net Service Name to identify the connection between Excel and Oracle Database.

> **Note:** The Spreadsheet Add-In uses features of Oracle Data Mining in Oracle Database Enterprise Edition. Whether the data to be mined is stored in Excel or in Oracle Database, an Oracle Net Service Name is required for accessing the data mining functionality.

The database connection used by the Spreadsheet Add-In must identify a database user that has been configured for data mining. Most of the privileges needed for data mining are also needed for predictive analytics. See "Creating Users" on page 2-7.

Remote Database

You can define the Net Service Name during installation of Oracle Client. "Installing Oracle Client" on page 3-4 illustrates the Easy Connect naming method.

Alternatively, you can use Oracle Net Configuration Assistant to create a Net Service Name. In either case, you will need to know the host name, the service name of the database instance, and the port number to create the Net Service Name. To use Oracle Net Configuration Assistant, follow the procedure described in "Local Database" and specify the host, service name, and port of the remote database service.

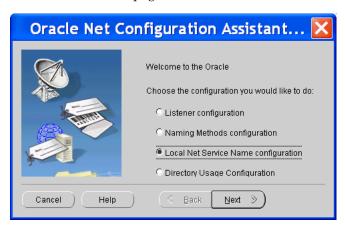
Local Database

You can install Oracle Database and Oracle Client on a single computer. You might do this if you wish to experiment with Oracle Data Mining on your personal computer or demonstrate Oracle Data Mining on a laptop.

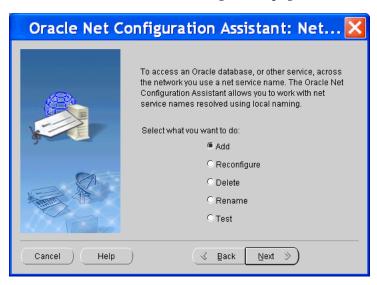
When the Oracle Database is local, you still need to create a Net Service Name for the Spreadsheet Add-In. Follow these steps on a Windows platform:

Open Oracle Net Configuration Assistant. Click **Start > Programs >** Oracle-- oracle_client_home > Configuration and Migration Tools > Net Configuration Assistant.

On the Welcome page, select Local Net Service Name configuration.



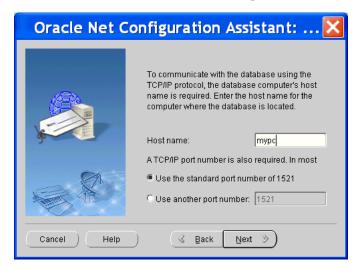
On the Net Service Name Configuration page, select **Add**.



On the Service Name page, type the service name of the database instance.



- On the Select Protocols page, select **TCP**.
- On the TCP/IP Protocol page, provide the name of the host computer where Oracle Database is installed and the port number.



- On the Test page, select **Yes**, **perform a test**.
- If the test fails, select Change Login and provide your own database user name and password. You will not see this page if the default SYSTEM user name and password can successfully make a connection.
- On the Net Service Name page, type a descriptive name to identify the connection to the database server.
 - You will choose this name when connecting to Oracle Database from Excel. Click Next.

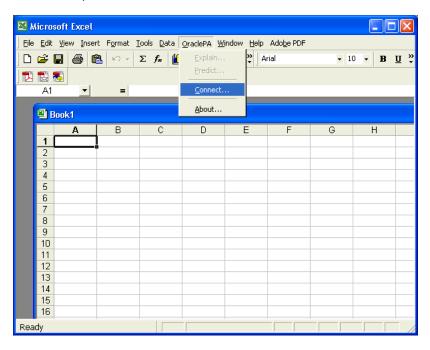


On the next page, choose **No** if you wish to exit the Net Configuration Assistant. Navigate through the final pages until you see the Finish button.

Connecting Excel to Oracle Database

Once you have created an Oracle Net Service Name for the instance of Oracle Database you plan to use, you can start using the Spreadsheet Add-In. You will specify the Net Service Name to connect the Add-In to Oracle Database. To make a connection, take these steps:

1. In Excel, click **OraclePA** > **Connect**.



On the Connect dialog box, select a service name from the drop-down list, and type in your database user name and password.

Installing and Using the Demo Programs

A number of demo programs are available with Oracle Data Mining. These programs illustrate the many features of the PL/SQL API, the Data Mining SQL functions, the Java API, and the BLAST table functions.

The demo programs create a set of models in the database. You can experiment with these models using either the APIs or Oracle Data Miner. You can examine the sample source code, which includes numerous comments, to familiarize yourself with the Oracle Data Mining APIs, and you can create your own models by modifying the samples.

This chapter includes the following sections:

- Requirements Checklist
- Obtaining the Demo Programs
- Creating a Demo User
- Examining the Data
- PL/SQL Demo Programs
- Java Demo Programs
- **Text Mining Demo Programs**
- **BLAST Demo Program**

Requirements Checklist

Several installation and configuration steps must be completed before you can run the Data Mining demo programs. Many of these steps are explained in this chapter. Some steps, such as installing Oracle Database, are explained in other chapters of this manual.

Complete these steps before attempting to run the Data Mining demo programs:

- Install Oracle Database 10g Release 2 (10.2) Enterprise Edition with the sample schemas, or obtain the connection information to an existing installation. See Chapter 1.
- **2.** Obtain the demo programs by installing the Database Companion or by downloading from the Oracle Technology Network web site. See "Obtaining the Demo Programs" on page 4-2.
- **3.** Create a user ID for running the demo programs. Grant the necessary privileges to this user. See "Creating a Demo User" on page 4-4 and "Granting Access Rights" on page 4-6.

- **4.** Populate the schema of the demo user with objects used by the demo programs. See "Populating the Schema" on page 4-7
- 5. To run the Java demos, check the requirements described in "Preparing to Run the Java Programs" on page 4-16.
- To run the BLAST demo, check the requirements described in "Preparing to Run the BLAST Demo" on page 4-20.

Obtaining the Demo Programs

The Oracle Data Mining sample programs are provided with Oracle Database Enterprise Edition. They are also available for download from the Oracle Technology Network.

Downloading from OTN

You can download a zip file containing the demo programs and documentation from the Oracle Data Mining page on OTN. Use the following URL.

http://www.oracle.com/technology/products/bi/odm/index.html

Extract the contents of the zip file, making sure to preserve the directory structure. The files will be copied to a directory called DATA MINING demos on your computer.

The zip file contains documentation similar to the information in this chapter. To access the documentation, open the file index.html in a browser.

Ensure that the SQL programs are located in a directory where you can access them using SQL*Plus. You may want to copy the Java programs to a Java development environment such as jDeveloper.

Installing Oracle Database Companion

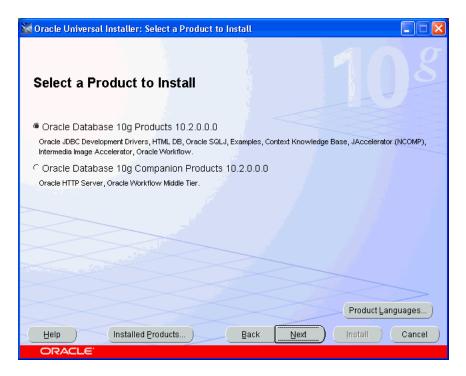
The Database Companion installation process copies the Oracle Data Mining demo programs, along with examples and demos of other database features, to the \RDBMS\demo subdirectory under ORACLE_HOME.

Note: In Oracle Database 10.1, Data Mining was installed in ORACLE HOME\dm, and the Data Mining demo programs were installed in ORACLE HOME\dm\demo\sample.

In Oracle 10g Release 2 (10.2), Data Mining is installed in the ORACLE HOME\RDBMS directory, and the Data Mining demo programs are installed in ORACLE HOME\RDBMS\demo.

To install the Database Companion on a Windows platform, take these steps:

- **1.** From the Companion installation directory, run SETUP. EXE. Oracle Universal Installer opens and displays the Welcome page. Click Next to advance to each page.
- 2. On the Select a Product to Install page, select **Oracle Database 10g Products** 10.2.0.0.0.

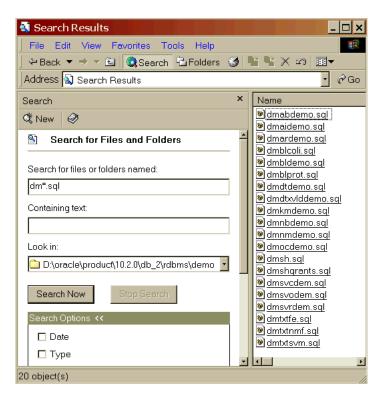


- On the Specify Home Details page, select the Oracle home in which you installed Oracle Database 10g Release 2. Do not rely on the default setting to be correct.
- On the Product-Specific Prerequisite Checks page, verify that all checks succeeded. If any checks failed, then you must correct the problem before proceeding.
- On the Summary page, review your previous choices, then click Install.
- On the End of Installation page, confirm that the installation was successful.

Verifying the Installation

You can find the SQL demos by searching for dm*.sql in ORACLE HOME\RDBMS\demo. You can find the Java demos by searching for dm*.java.

On a Windows platform, you can use Windows File Manager to list the SQL programs as follows.



You can use Windows File Manager to list the Java programs as follows.



Creating a Demo User

Data Mining demo users require several database permissions, as well as SELECT access to tables in the SH sample schema. Users that will primarily use the demo programs, and not be mining large data sets, do not need personal tablespaces.

Note: In Oracle Database 10.1, a data mining user account, called DMUSER, was provided with the software. In Oracle Database 10g Release 2 (10.2), DMUSER is no longer provided; you must create your own data mining user accounts.

To create a demo user for Oracle Data Mining, take these steps:

- Create a user name and password.
- Run the dmshgrants script to grant the necessary privileges to the user.
- Run the dmsh script to populate the user's schema with objects that support the demo programs.

Using SQL*Plus to Create the User

To create the user in SQL*Plus, log in as the SYSTEM user and type a command like the following:

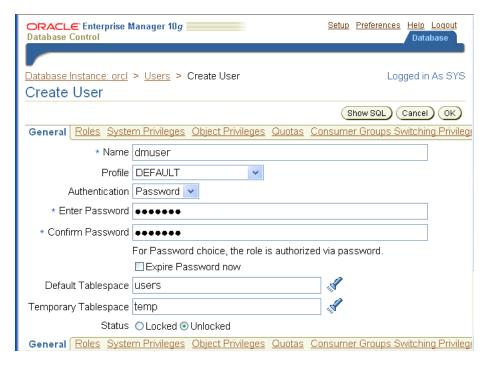
```
CREATE USER dmuser IDENTIFIED BY dmpsw
    DEFAULT TABLESPACE users
    TEMPORARY TABLESPACE temp
    QUOTA UNLIMITED on users;
```

This command creates the user dmuser with the password dmpsw. It provides default access to two tablespaces shared by several other sample schemas.

Using Enterprise Manager to Create the User

To create the user in Enterprise Manager, follow steps like these:

- Open Enterprise Manager and log in as the SYSTEM user.
- **2.** On the Administration page, choose **Users**.
- On the Users page, click the **Create** button.
- **4.** On the Create Users page, provide information like this.



You can use the Quota page to grant unlimited quota to the USERS schema.

Granting Access Rights

The dmshgrants script grants database privileges to a demo user. These privileges are the same as those listed in "Access Rights" on page 2-7.

The dmshgrants script is provided by the Database Companion installation. You can locate it in the \rdbms\demo subdirectory of ORACLE HOME. The dmshgrants script is also included with the Data Mining demos in the zip file you can download from OTN.

In addition to granting database privileges, dmshqrants grants SELECT rights on these tables in the SH schema:

```
COUNTRIES
CUSTOMERS
PRODUCTS
SUPPLEMENTARY DEMOGRAPHICS
SALES
```

For text mining, dmshgrants grants access rights to an Oracle Text package:

```
EXECUTE ON ctxsys.ctx ddl
```

If you do not have access to dmshgrants, you can set permissions using the SQL GRANT command or Enterprise Manager Database Control.

Run DMSHGRANTS

The following command, executed in SQL*Plus on a Windows platform, passes two arguments to dmshgrants: The password for the SH user (shpsw in this case) and the user name receiving the access rights (dmuser).

@%ORACLE HOME%\rdbms\demo\dmshgrants shpsw dmuser

On Linux, this is the equivalent command:

@\$ORACLE HOME/rdbms/demo/dmshgrants shpsw dmuser

Populating the Schema

The dmsh script populates the schema of the demo user with tables and views that reference the source data in the SH schema. It also creates tables and indexes that support text mining.

The dmsh script, like dmshgrants, is provided by the Database Companion installation, and is included with the Data Mining demos in the zip file you can download from OTN.

Run DMSH

To run dmsh, log in to SQL*Plus as the data mining user (in this case, dmuser with the password dmpsw).

```
SOL>Connect dmuser/dmpsw
SQL>@%ORACLE HOME%\rdbms\demo\dmsh
SQL>commit;
```

Sample Schemas

The Oracle Data Mining demo programs use data in the SH sample schema. If the sample schemas were not originally installed in the database, you must install SH.

To install the sample schemas after installation of Oracle Database, use Database Configuration Assistant. For an example that uses Database Configuration Assistant, see "Adding the Data Mining Option to a Database" on page 1-9.

To install SH without the other sample schemas, run the sh main script. The sh main.sql file is located in \demo\schema\sales history in the Oracle home directory. You must run the script as SYS.

See Also: Oracle Database Sample Schemas for more details about installing the sample schemas.

A Sample User Scenario

Let's assume that you have access to Oracle Database 10g Release 2 (10.2), which has been installed on a Windows host with the Data Mining option, the sample schemas, and the sample programs. You can log in to this database as SYS using SQL*Plus.

Several Data Mining users (dmuser1 and dmuser2) have already been created in this database. You want to create your own user ID (dmuser3) and run the demo programs to create some starter models in your schema. To accomplish this, you would perform the following steps.

1. Log in to the database as SYS and create the dmuser3 user.

```
Enter user-name: sys / as sysdba
   Enter password: sys password
SQL> create user dmuser3 identified by dmuser3 password
    default tablespace users temporary tablespace temp
     quota unlimited on users;
```

2. Run dmshgrants.sql to grant privileges to dmuser3.

```
SQL> @ %ORACLE_HOME%\rdbms\demo\dmshgrants SH_password dmuser3
```

3. Connect as dmuser3 and run dmsh.sql to populate the dmuser3 schema with views and tables needed by the demo programs. Save your changes in the database.

```
SQL> connect dmuser3/dmuser3 password
SQL> @ %ORACLE HOME%\rdbms\demo\dmsh
SQL> commit;
```

You can now run any of the PL/SQL data mining demos. For example, while logged in to SQL*Plus as dmuser3, you could run the Naive Bayes PL/SQL demo with the following statement.

```
SQL>@ %ORACLE HOME%\rdbms\demo\dmnbdemo
```

You can also run any of the Java data mining demos, if you have Java 1.4.2 or higher and your CLASSPATH is set as described in "Preparing to Run the Java Programs" on page 4-16. For example, you could run the Naive Bayes Java demo with a command like the following at the operating system prompt.

```
>java dmnbdemo myserver:1521:orcl dmuser3 dmuser3_password
```

Note that the BLAST demo uses different data sets and requires a separate setup procedure. See "Preparing to Run the BLAST Demo" on page 4-20.

> **Note:** All the demo programs are re-executable. They start by deleting the results of the previous run before executing the current run.

Examining the Data

The dmsh.sql script creates views, tables, and indexes in the user's schema. The views define columns of customer data from tables in the SH schema. This data is used by the Data Mining demo programs. The tables reference the same columns in SH, but they include an extra COMMENTS column for text mining. The indexes are used to extract terms from the text in the COMMENTS column and build a nested table column.

Customer Data for Data Mining

Views in the data mining user's schema define columns of data from the CUSTOMERS, SALES, PRODUCTS, COUNTRIES, and SUPPLEMENTARY DEMOGRAPHICS tables in the SH schema. You can list these views with the following SQL statements.

```
SQL>connect dmuser3/dmuser3 password
SQL>select view name from user views;
```

The views are listed in Table 4–1.

Table 4–1 Views Used by the Data Mining Demo Programs

View Name	Description
MINING_DATA_APPLY_STR_V	Scoring data for o-cluster
MINING_DATA_BUILD_STR_V	Build data for o-cluster
MINING_DATA_APPLY_V	Scoring data for data mining (not text mining)
MINING_DATA_BUILD_V	Build data for data mining (not text mining)

Table 4-1 (Cont.) Views Used by the Data Mining Demo Programs

View Name	Description
MINING_DATA_TEST_V	Test data for data mining (not text mining)
MARKET_BASKET_V	Data for association rules
MINING_DATA_ONE_CLASS_V	Data for one-class SVM

You can see the references to tables in SH by listing the view definitions. The definition of the view MINING_DATA_BUILD_V is shown as follows.

```
SQL>set long 1000000
SQL>set longc 100000
SQL>select text from all views where
   owner='DMUSER3'and view_name='MINING_DATA_BUILD_V';
      SELECT a.CUST ID, a.CUST GENDER, 2003-a.CUST YEAR OF BIRTH AGE,
            a.CUST_MARITAL_STATUS, c.COUNTRY_NAME, a.CUST_INCOME_LEVEL,
            b.EDUCATION, b.OCC UPATION, b.HOUSEHOLD SIZE, b.YRS RESIDENCE,
            b.AFFINITY CARD, b.BULK PACK DISKETTES, b.FLAT PANEL MONITOR,
            b.HOME_THEATER_PACKAGE, b.BOOKKEEPING _APPLICATION,
            b.PRINTER_SUPPLIES, b.Y_BOX_GAMES, b.OS_DOC_SET_KANJI
       FROM sh.customers a,
            sh.supplementary_demographics b,
            sh.countries c
       WHERE a.CUST ID = b.CUST ID AND a.country id = c.country id
            AND a.cust_id between 101501 and 103000
```

The views are used to build, test, and score the sample models. Each view has a CUSTOMER ID column, which is the case ID, and an AFFINITY CARD column, which is the target used by the predictive models. Most of the views provide data for 1500 customers (1500 rows). The view used by the one-class SVM model has data for 940 customers.

The columns of build data in the MINING DATA BUILD V view are listed in the following example.

SQL> describe MINING DATA BUILD V

CUST_ID	NOT NULL	NUMBER
CUST_GENDER	NOT NULL	CHAR(1)
AGE		NUMBER
CUST_MARITAL_STATUS		VARCHAR2 (20)
COUNTRY_NAME	NOT NULL	VARCHAR2 (40)
CUST_INCOME_LEVEL		VARCHAR2(30)
EDUCATION		VARCHAR2 (21)
OCCUPATION		VARCHAR2 (21)
HOUSEHOLD_SIZE		VARCHAR2 (21)
YRS_RESIDENCE		NUMBER
AFFINITY_CARD		NUMBER(10)
BULK_PACK_DISKETTES		NUMBER(10)
FLAT_PANEL_MONITOR		NUMBER(10)
HOME_THEATER_PACKAGE		NUMBER(10)
BOOKKEEPING_APPLICATION		NUMBER(10)
PRINTER_SUPPLIES		NUMBER(10)
Y_BOX_GAMES		NUMBER(10)
OS_DOC_SET_KANJI		NUMBER(10)

Market Basket Data for Association Rules

The association rules demos use the MARKET BASKET V data set, which includes columns of products from the PRODUCTS table and the CUSTOMER ID column from the CUSTOMERS table in SH. The columns of the MARKET BASKET V view are listed in the following example.

SQL> describe MARKET_BASKET_V

CUST_ID	NOT	NULL	NUMBER
EXTENSION_CABLE			NUMBER
FLAT_PANEL_MONITOR			NUMBER
CD_RW_HIGH_SPEED_5_PACK			NUMBER
ENVOY_256MB_40GB			NUMBER
ENVOY_AMBASSADOR			NUMBER
EXTERNAL_8X_CD_ROM			NUMBER
KEYBOARD_WRIST_REST			NUMBER
SM26273_BLACK_INK_CARTRIDGE			NUMBER
MOUSE_PAD			NUMBER
MULTIMEDIA_SPEAKERS_3INCH			NUMBER
OS_DOC_SET_ENGLISH			NUMBER
SIMM_16MB_PCMCIAII_CARD			NUMBER
STANDARD_MOUSE			NUMBER

Customer Data for Text Mining

The text mining demos use the same customer data from tables in SH, but they include either an extra text column or a collection type column. The collection type is a nested table of type DM_NESTED_NUMERICALS.

You can list these tables with the following SQL statements.

```
SQL>connect dmuser3/dmuser3_password
SQL>select table name from user tables where table name like '%MINING%';
```

The tables are listed in Table 4–2.

Table 4–2 Tables Used by the Text Mining Demo Programs

Table Name	Description
MINING_APPLY_NESTED_TEXT	Apply table with COMMENTS column as DM_NESTED_NUMERICALS
MINING_BUILD_NESTED_TEXT	Build table with COMMENTS column as DM_NESTED_NUMERICALS
MINING_TEST_NESTED_TEXT	Test table with COMMENTS column as DM_NESTED_NUMERICALS
MINING_APPLY_TEXT	Apply table with COMMENTS column as VARCHAR2 (4000)
MINING_BUILD_TEXT	Build table with COMMENTS column as VARCHAR2 (4000)
MINING_TEST_TEXT	Test table with COMMENTS column as VARCHAR2 (4000)

In the MINING BUILD TEXT, MINING TEST TEXT, and MINING APPLY TEXT tables, the COMMENTS column is of type VARCHAR2 (4000).

```
SQL> describe MINING BUILD TEXT
Name
                                  Null?
                                           Type
 ______
CUST ID
                                   NOT NULL NUMBER
CUST GENDER
                                   NOT NULL CHAR (1)
AGE
                                          NUMBER
CUST MARITAL STATUS
                                          VARCHAR2 (20)
                                 NOT NULL VARCHAR2 (40)
COUNTRY_NAME
```

CUST INCOME LEVEL	VARCHAR2(30)
COST_INCOME_DEVED	, , ,
EDUCATION	VARCHAR2(21)
OCCUPATION	VARCHAR2(21)
HOUSEHOLD_SIZE	VARCHAR2(21)
YRS_RESIDENCE	NUMBER
AFFINITY_CARD	NUMBER(10)
BULK_PACK_DISKETTES	NUMBER(10)
FLAT_PANEL_MONITOR	NUMBER(10)
HOME_THEATER_PACKAGE	NUMBER(10)
BOOKKEEPING_APPLICATION	NUMBER(10)
PRINTER_SUPPLIES	NUMBER(10)
Y_BOX_GAMES	NUMBER(10)
OS_DOC_SET_KANJI	NUMBER(10)
COMMENTS	VARCHAR2 (4000)

In the MINING \star NESTED TEXT tables, the COMMENTS column is of type DM_ NESTED_NUMERICALS.

SQL> describe MINING_BUILD_NESTED_TEXT			
Name	Nul	1?	Туре
CUST_ID			NUMBER
CUST_GENDER	NOT	NULL	CHAR(1)
AGE			NUMBER
CUST_MARITAL_STATUS			VARCHAR2 (20)
COUNTRY_NAME	NOT	${\tt NULL}$	VARCHAR2 (40)
CUST_INCOME_LEVEL			VARCHAR2(30)
EDUCATION			VARCHAR2(21)
OCCUPATION			VARCHAR2 (21)
HOUSEHOLD_SIZE			VARCHAR2 (21)
YRS_RESIDENCE			NUMBER
AFFINITY_CARD			NUMBER(10)
BULK_PACK_DISKETTES			NUMBER(10)
FLAT_PANEL_MONITOR			NUMBER(10)
HOME_THEATER_PACKAGE			NUMBER(10)
BOOKKEEPING_APPLICATION			NUMBER(10)
PRINTER_SUPPLIES			NUMBER(10)
Y_BOX_GAMES			NUMBER(10)
OS_DOC_SET_KANJI			NUMBER(10)
COMMENTS			DMSYS.DM_NESTED_NUMERICALS

The process of extracting terms from a text column into a nested table column is described in "Text Mining Demo Programs" on page 4-17 and in the Oracle Data Mining Application Developer's Guide.

PL/SQL Demo Programs

The PL/SQL demo programs illustrate the use of the DBMS DATA MINING package for creating models and the DBMS DATA MINING TRANSFORM package for performing transformations on the mining data.

Table 4–3 lists the PL/SQL demo programs.

Table 4–3 Mining Functions in PL/SQL demos

Mining Function	Algorithm	Program File
Association Rules	Apriori	dmardemo.sql
Attribute Importance	Minimum Descriptor Length	dmaidemo.sql

Table 4–3 (Cont.) Mining Functions in PL/SQL demos

Mining Function	Algorithm	Program File
Classification	Adaptive Bayes Network	dmabdemo.sql
Classification	Decision Tree	dmdtdemo.sql
Classification	Decision Tree (cross validation)	dmdtxvlddemo.sql
Classification	Naive Bayes	dmnbdemo.sql
Classification	Support Vector Machines	dmsvcdem.sql
Classification	Support Vector Machines (One Class)	dmsvodem.sql
Clustering	k-Means	dmkmdemo.sql
Clustering	O-Cluster	dmocdemo.sql
Feature Extraction	Non-Negative Matrix Factorization	dmnmdemo.sql
Regression	Support Vector Machines	dmsvrdem.sql
Text Mining	Term extraction using CTX procedures	dmtxtfe.sql
Text Mining	Non-Negative Matrix Factorization	dmtxtnmf.sql
Text Mining	Support Vector Machines (Classification)	dmtxtsvm.sql

Note: The PL/SQL text mining demos are described in "Text Mining" in PL/SQL" on page 4-18.

See Also: Oracle Database PL/SQL Packages and Types Reference and Oracle Data Mining Application Developer's Guide for information on the Oracle Data Mining PL/SQL API. See Oracle Database SQL Reference for information on the SQL functions for data mining.

PL/SQL Program Summaries

Summary descriptions of the PL/SQL demo programs are provided in Table 4–4. For detailed descriptions of the demo programs, see the comments in the source code.

Table 4–4 Overview of the PL/SQL Demo Programs

Mining Function	Description
Classification	The classification programs demonstrate various preprocessing techniques and perform the following steps:
	 Build a classification model using training data
	 Display model details and settings
	 Test the model by applying the model on the test data
	 Present test metrics, such as confusion matrix, lift, and ROC
	 Apply the model on the scoring data
	 Present apply results
	 Present ranked apply results, influenced by a cost matrix
	The dmdtxvlddemo.sql program demonstrates cross-validation techniques for decision tree based-classification. With minor modifications, this program can be used to perform cross validation using other models/algorithms.

Table 4-4 (Cont.) Overview of the PL/SQL Demo Programs

Mining Function	Description
Regression	dmsvrdem.sql uses different test metrics, but otherwise performs most of the same steps used in the classification programs. Selected attributes of the input data are preprocessed (normalized).
Association	dmardemo.sql builds an association model and presents frequent itemsets and association rules as output. Selected attributes of the input data are preprocessed (binned).
Clustering	dmkmdemo.sql (k-Means) and dmocdemo.sql (0-cluster) build clustering models and present cluster details, such as rules, centroid, and histogram for each cluster as output. The models are scored, and the probabilities associated with each cluster are returned as output. Selected attributes of the input data are preprocessed (normalized).
Feature Extraction	dmnmdemo.sql builds a feature extraction model and presents model details as the output. The model is scored, and each feature ID is associated with a probability. Selected attributes of the input data are preprocessed (normalized).
Attribute Importance	dmaidemo.sql builds an attribute importance model and presents a list of important attributes as the output of model details. Selected attributes of the input data are preprocessed (binned).

Data Mining SQL Functions

Note: The SQL functions for Data Mining are documented in *Oracle* Database SQL Reference. Information about these functions is also provided in Oracle Data Mining Application Developer's Guide, where they are referred to as *SQL* scoring functions.

Some of the PL/SQL demo programs use Data Mining SQL functions to apply models created with the DBMS_DATA_MINING package. The Data Mining functions can also be used to apply models created with the Java API.

The programs that demonstrate the Data Mining functions are listed in Table 4–5.

Table 4–5 Data Mining SQL Functions in the Demo Programs

Program Name	Algorithm	SQL Functions Used
dmkmdemo.sql	k-Means	CLUSTER_ID CLUSTER_PROBABILITY CLUSTER_SET
dmocdemo.sql	O-Cluster	CLUSTER_ID
dmnmdemo.sql	NMF	FEATURE_ID FEATURE_SET FEATURE_VALUE
dmdtdemo.sql	Decision Tree	PREDICTION PREDICTION_COST PREDICTION_DETAILS PREDICTION_SET
dmsvcdem.sql	SVM Classification	PREDICTION PREDICTION_PROBABILITY PREDICTION_SET
dmsvodem.sql	SVM One Class	PREDICTION PREDICTION_PROBABILITY

Table 4-5 (Cont.) Data Mining SQL Functions in the Demo Programs

Program Name	Algorithm	SQL Functions Used
dmsvrdem.sql	SVM Regression	PREDICTION
dmtxtsvm.sql	Text Mining	PREDICTION PREDICTION_PROBABILITY

Running the PL/SQL Programs

In SQL*Plus, use commands like the following to execute the sample programs and list the models created by them.

```
>sqlplus dmuser3/dmuser3_password
SQL> set serveroutput on
SQL> set echo on
SQL> @ %ORACLE_HOME%\rdbms\demo\program_name
SQL> set linesize 200
SQL> set pagesize 100
SQL> select NAME, FUNCTION NAME, ALGORITHM NAME, TARGET_ATTRIBUTE from DM_USER_MODELS;
```

NAME	FUNCTION_NAME	ALGORITHM_NAME	TARGET_ATTRIBUTE
T_NMF_SAMPLE	FEATURE_EXTRACTION	NONNEGATIVE_MATRIX_FACTOR	
T_SVM_CLAS_SAMPLE	CLASSIFICATION	SUPPORT_VECTOR_MACHINES	AFFINITY_CARD
AR_SH_SAMPLE	ASSOCIATION_RULES	APRIORI_ASSOCIATION_RULES	
AI_SH_SAMPLE	ATTRIBUTE_IMPORTANCE	MINIMUM_DESCRIPTION_LENGTH	AFFINITY_CARD
ABN_SH_CLAS_SAMPLE	CLASSIFICATION	ADAPTIVE_BAYES_NETWORK	AFFINITY_CARD
DT_SH_CLAS_SAMPLE	CLASSIFICATION	DECISION_TREE	AFFINITY_CARD
NB_SH_CLAS_SAMPLE	CLASSIFICATION	NAIVE_BAYES	AFFINITY_CARD
SVMC_SH_CLAS_SAMPLE	CLASSIFICATION	SUPPORT_VECTOR_MACHINES	AFFINITY_CARD
OC_SH_CLUS_SAMPLE	CLUSTERING	O_CLUSTER	
KM SH CLUS SAMPLE	CLUSTERING	KMEANS	
NMF_SH_SAMPLE	FEATURE_EXTRACTION	NONNEGATIVE_MATRIX_FACTOR	
SVMR_SH_REGR_SAMPLE	REGRESSION	SUPPORT_VECTOR_MACHINES	AGE

Java Demo Programs

The Java demos illustrate the features of the Oracle Data Mining Java API, which implements Oracle-specific extensions to the Java Data Mining (JDM) 1.0 standard.

The Java demos demonstrate data preprocessing and the basic mining functions. Additional Java samples demonstrate predictive analytics, import/export, and text mining.

Table 4–6 lists the Java programs that illustrate the basic mining functions.

Table 4–6 Mining Functions in Java Demos

Mining Function	Algorithm	Program File
Association Rules	Apriori	dmardemo.java
Attribute Importance	Minimum Descriptor Length	dmaidemo.java
Classification	Adaptive Bayes Network	dmabdemo.java
Classification	Decision Trees	dmtreedemo.java

Table 4-6 (Cont.) Mining Functions in Java Demos

Mining Function	Algorithm	Program File
Classification	Naive Bayes	dmnbdemo.java
Classification	Support Vector Machines	dmsvcdemo.java
Classification	Support Vector Machines (one class)	dmsvodemo.java
Clustering	K-Means	dmkmdemo.java
Clustering	O-Cluster	dmocdemo.java
Feature Extraction	Non-Negative Matrix Factorization	dmnmdemo.java
Regression	Support Vector Machines	dmsvrdemo.java
Text Mining	Non-Negative Matrix Factorization	dmtxtnmfdemo.java
Text Mining	Support Vector Machines (Classification)	dmtxtsvmdemo.java

Table 4–7 lists the Java programs that illustrate special mining tasks. These features are all supported in the PL/SQL API as well, since the Java API is layered on the PL/SQL API.

Table 4–7 Mining Tasks in Java Demos

Mining Task	Description	Demo Program File
Data Transformations	Binning, clipping, and normalization	dmxfdemo.java
Predictive Analytics	Automated predict and explain	dmpademo.java
Model Export/Import	To/from Data Pump dump file	dmexpimpdemo.java
Classification Model Scoring	Ways of applying an NB model	dmapplydemo.java

Note: The Java text mining demos are described in "Text Mining in Java" on page 4-19.

See Also: Oracle Data Mining Java API Reference (javadoc) and the Oracle Data Mining Application Developer's Guide for information on the Java API.

Java Program Summaries

Summary descriptions of the Java demo programs are provided in Table 4–8. For detailed descriptions, see the comments in the source code.

Table 4-8 Overview of the Java Demo Programs

Mining Function or Task	Description
Classification	The classification programs demonstrate various preprocessing techniques and perform the following steps:
	 Build a classification model using training data
	 Display model details and settings
	 Test the model by applying the model on the test data
	■ Present test metrics, such as confusion matrix, lift, and ROC
	 Apply the model on the scoring data
	 Present apply results
	 Present ranked apply results, influenced by a cost matrix
	The dmapplydemo.java program demonstrates several ways of applying a Naive Bayes model.
Regression	dmsvrdemo.java uses different test metrics, but otherwise performs most of the same steps used in the classification programs. Selected attributes of the input data are preprocessed (normalized).
Association	dmardemo.java builds an association model and presents frequent itemsets and association rules as output. Selected attributes of the input data are preprocessed (binned).
Clustering	dmkmdemo.java (k-Means) and dmocdemo.java (0-cluster) build clustering models and present cluster details, such as rules, centroid, and histogram for each cluster as output. The models are scored, and the probabilities associated with each cluster are returned as output. Selected attributes of the input data are preprocessed (normalized).
Feature Extraction	dmnmdemo.java builds a feature extraction model and presents model details as the output. The model is scored, and each feature ID is associated with a probability. Selected attributes of the input data are preprocessed (normalized).
Attribute Importance	dmaidemo.java builds an attribute importance model and presents a list of important attributes as the output of model details. Selected attributes of the input data are preprocessed (binned).
Data Transformations	dmxfdemo.java demonstrates binning, clipping, and normalization transformations.
Predictive Analytics	${\tt dmpademo}.{\tt java}{\tt demonstrates}$ PREDICT and EXPLAIN functions.
Model Import/Export	dmexpimpdemo.java builds a Naive Bayes model, exports it to a dump file, then imports it from the dump file.

Preparing to Run the Java Programs

Before running the Java programs, do the following:

1. Check that the version of Java you are using is 1.4.2. You can execute the following in a command window to check the version of Java:

```
>java -version
```

2. Ensure that your CLASSPATH includes the following JAR files:

```
%ORACLE_HOME%\rdbms\jlib\jdm.jar
%ORACLE_HOME%\rdbms\jlib\ojdm_api.jar
```

```
%ORACLE HOME%\rdbms\jlib\xdb.jar
%ORACLE_HOME%\jdbc\lib\ojdbc14.jar
%ORACLE HOME%\oc4j\j2ee\home\lib\connector.jar
%ORACLE HOME%\jlib\orai18n.jar
%ORACLE HOME%\jlib\orai18n-mapping.jar
%ORACLE HOME%\lib\xmlparserv2.jar
```

3. Compile the programs. Use the javac executable in the \jdk\bin\ directory within Oracle home.

```
>javac program_name
```

Running the Java Programs

Use the following syntax to execute the Java demo programs. Use the java executable in the \jdk\bin\ directory within Oracle home.

```
>java program_name connect_string user_name user_password
```

The connection string specifies your Oracle database connection. It identifies the machine hosting the database, the port through which the connection is made, and the name of the database instance (the Oracle system identifier).

```
host name:port:SID
```

For example, the following command executes the Naive Bayes demo dmnbdemo.java as dmuser3 in the database instance orcl on host mach05 at port 1521.

```
>java dmnbdemo mach05:1521:orcl dmuser3 dmuser3 password
```

You can list the models created by the Java programs with commands like the following in SQL*Plus.

```
>sqlplus dmuser3/dmuser3 password
SOL> set linesize 200
SQL> set pagesize 100
SQL> select NAME, FUNCTION_NAME, ALGORITHM_NAME, TARGET_ATTRIBUTE
           from dm_user_models where NAME like '%JDM';
```

NAME	FUNCTION_NAME	ALGORITHM_NAME	TARGET_ATTRIBUTE
TXTNMFMODEL_JDM	FEATURE_EXTRACTION	NONNEGATIVE_MATRIX_FACTOR	
ARMODEL_JDM	ASSOCIATION_RULES	APRIORI_ASSOCIATION_RULES	
TREEMODEL_JDM	CLASSIFICATION	DECISION_TREE	AFFINITY_CARD
AIMODEL_JDM	ATTRIBUTE_IMPORTANCE	MINIMUM_DESCRIPTION_LENGTH	AFFINITY_CARD
ABNMODEL_JDM	CLASSIFICATION	ADAPTIVE_BAYES_NETWORK	AFFINITY_CARD
NBMODEL_JDM	CLASSIFICATION	NAIVE_BAYES	AFFINITY_CARD
SVMCMODEL_JDM	CLASSIFICATION	SUPPORT_VECTOR_MACHINES	AFFINITY_CARD
SVMOMODEL_JDM	CLASSIFICATION	SUPPORT_VECTOR_MACHINES	
KMMODEL_JDM	CLUSTERING	KMEANS	
OCMODEL_JDM	CLUSTERING	O_CLUSTER	
NMFMODEL_JDM	FEATURE_EXTRACTION	NONNEGATIVE_MATRIX_FACTOR	
SVMRMODEL_JDM	REGRESSION	SUPPORT_VECTOR_MACHINES	AGE
TXTSVMMODEL_JDM	CLASSIFICATION	SUPPORT_VECTOR_MACHINES	AFFINITY_CARD

Text Mining Demo Programs

Oracle Data Mining can mine text columns that have undergone pre-processing by Oracle Text routines.

Oracle Text is a technology for building text query and document classification applications. It provides indexing, word and theme searching, and viewing capabilities for text. Oracle Text is included in a general installation of Oracle Database Enterprise Edition, and therefore is already present in a database installed according to the instructions in Chapter 1.

The pre-processing steps for text mining create nested table columns of type DM NESTED NUMERICALS from columns of type VARCHAR2 or CLOB. Each row of the nested table specifies an attribute name and a value. The type definition is as follows.

```
CREATE OR REPLACE TYPE DM NESTED NUMERICAL AS OBJECT
  (attribute_name VARCHAR2(30),
            NUMBER)
  value
CREATE OR REPLACE TYPE DM NESTED NUMERICALS AS TABLE OF DM NESTED NUMERICAL
```

Terms extracted from text documents into nested tables can become generic attributes in training or scoring data. Classification, clustering, and feature-extraction models can be built using these attributes.

Sample text mining programs in both PL/SQL and Java illustrate classification and feature extraction of a pre-processed text column.

Text Mining in PL/SQL

Three PL/SQL demo programs illustrate the process of text mining. One program illustrates the pre-processing that is required to prepare the data for mining. The other two programs build models that use the transformed text.

Text Transformation Demo

To prepare a column for text mining using the PL/SQL API, you must use Oracle Text routines to perform the following general steps:

- **1.** Create a domain index on the column.
- Use the index to extract terms from the column to a temporary table.
- 3. Populate a column of type DM NESTED NUMERICALS with the terms in the temporary table.

The process of term extraction using Oracle Text is illustrated in the demo program dmtxtfe.sql. The source code contains extensive comments that explain the steps involved in transforming text into a set of features that can be mined using Oracle Data Mining.

More details about text transformation are provided in the *Oracle Data Mining* Application Developer's Guide.

Text Transformation for the PL/SQL Text Mining Demo Programs

The dmsh.sql script performs the text transformation required by the PL/SQL text mining demos. There are two such demos: dmtxtnmf.sql, which builds a feature extraction model using Non-Negative Matrix Factorization, and dmtxtsvm.sql, which builds a classification model using Support Vector Machines. Both of these programs use the tables that have a nested table column of comment data.

```
MINING BUILD NESTED TEXT
MINING TEST NESTED TEXT
MINING APPLY NESTED TEXT
```

The Sample Text Mining Models (PL/SQL)

You can run the PL/SQL text mining demo programs, dmtxtnmf.sql and dmtxtsvm.sql, like the other PL/SQL programs. The models created by these programs are listed in the following example.

```
SQL> @ %ORACLE HOME%\rdbms\demo\dmtxtnmf.sql
SQL> @ %ORACLE HOME%\rdbms\demo\dmtxtsvm.sql
SQL> select NAME, FUNCTION NAME, ALGORITHM NAME, TARGET ATTRIBUTE
           from dm_user_models;
```

NAME	FUNCTION_NAME	ALGORITHM_NAME	TARGET_ATTRIBUTE
T_NMF_SAMPLE	FEATURE_EXTRACTION	NONNEGATIVE_MATRIX_FACTOR	
T SVM CLAS SAMPLE	CLASSIFICATION	SUPPORT VECTOR MACHINES	AFFINITY CARD

Text Mining in Java

Two Java demo programs illustrate the process of text mining. One builds a feature extraction model, the other builds a classification model.

Text Transformation for the Java Demo Text Mining Programs

The Oracle Data Mining Java API provides an interface that handles the term extraction process. If you are developing data mining applications in Java, you do not need to use Oracle Text directly. However, you must ensure that Oracle Text is present in the database.

The OraTextTransform interface is used to perform text transformation within the Java text mining demos. There are two such demos: dmtxtnmfdemo.java, which builds a feature extraction model using Non-Negative Matrix Factorization, and dmtxtsvmdemo. java, which builds a classification model using Support Vector Machines. Both of these programs create build, test, and apply data sets from the tables that have a text column of comment data.

```
MINING BUILD TEXT
MINING TEST TEXT
MINING APPLY TEXT
```

The Sample Text Mining Models (Java)

You can run the Java text mining demo programs, dmtxtnmfdemo.java and dmtxtsvmdemo.java, like the other Java programs. The models created by these programs are shown in the following example.

```
> java dmtxtnmfdemo host:port:SID dmuser3 dmuser3_password
> java dmtxtsvmdemo host:port:SID dmuser3 dmuser3 password
>sqlplus dmuser3/dmuser3 password
SQL> select NAME, FUNCTION NAME, ALGORITHM NAME, TARGET ATTRIBUTE
           from dm user models;
```

NAME	FUNCTION_NAME	ALGORITHM_NAME	TARGET_ATTRIBUTE
txtnmfModel_jdm	FEATURE_EXTRACTION	NONNEGATIVE_MATRIX_FACTOR	
txtsvmModel_jdm	CLASSIFICATION	SUPPORT_VECTOR_MACHINES	AFFINITY_CARD

BLAST Demo Program

The Oracle implementation of the Basic Local Alignment Search Tool (BLAST) is demonstrated in the demo program, dmbldemo.sql. This program provides examples of sequence matching queries using the BLAST table functions.

> **Note:** The BLAST algorithm detects local alignments in nucleotide and protein databases. BLAST performs a kind of data mining, since it finds regions of similarity embedded in otherwise unrelated sequences. However, BLAST does not use Oracle Data Mining technology. It is a completely independent tool.

> For more information about the Oracle implementation of BLAST, see Oracle Data Mining Application Developer's Guide

The BLAST demo program and configuration scripts are listed in Table 4–9.

Table 4-9 BLAST Demo Files

File Name	Description
dmbldemo.sql	BLAST demo program.
dmblcoli.sql	Script that creates the Ecoli data set.
dmblprot.sql	Script that creates the Swiss Protein data set.
dmblprot.txt	Text for SQL*Loader to load into the Swiss Protein data set.
dmbl.ctl	Control file for SQL*Loader to use in loading the Swiss Protein data set.

See Also: Oracle Data Mining Application Developer's Guide and Oracle Data Mining Concepts for information on BLAST.

Preparing to Run the BLAST Demo

The demo BLAST table functions in dmbldemo.sql use two data sets: SwissProt and ecoli10. To prepare these data sets, log in to SQL*Plus as the data mining user and run the dmblprot and dmblcoli scripts as shown in the following example.

```
SQL>connect dmuser3/dmuser3 password
SQL>@ %ORACLE HOME%\rdbms\demo\dmblprot.sql
SQL>@ %ORACLE_HOME%\rdbms\demo\dmblcoli.sql
```

Exit SQL*Plus and use the SQL*Loader utility to load data into the SwissProt database in the schema of the data mining user. From the command prompt, change to the \rdbms\demo directory under Oracle home and execute the following command.

```
>sqlldir dmuser3/dmuser3 password
        control=dmblprot.ctl data=dmblprot.txt log=dmblprot.log
```

Running the BLAST Table Functions

The demo program dmbldemo.sql contains multiple invocations of BLAST table functions. You can run them all at once by running the dmbldemo.sql script, or you can copy individual table functions to the SQL*Plus command line and execute them individually.

To run the demo program, log in to SQL*Plus as the data mining user and run the dmbldemo.sql script as shown in the following example.

```
SQL>connect dmuser3/dmuser3 password
SQL> @ %ORACLE_HOME%\rdbms\demo\dmbldemo.sql
```

The following example shows how you could execute the sample BLASTP_MATCH table function at the SQL*Plus command line. BLASTP_MATCH is run against all the human proteins in the SwissProt database.

```
SQL>connect dmuser3/dmuser3 password
SQL>SET serveroutput ON
SQL>SET trimspool ON
SQL>SET pages 10000
SQL>column seq_id format a10
SQL>select T_SEQ_ID AS seq_id, score, EXPECT as evalue
         from TABLE(
              BLASTP MATCH (
                  (select sequence from query db),
                    CURSOR(SELECT seq_id, seq_data
                          FROM swissprot
                           WHERE organism = 'Homo sapiens (Human)'),
                   1,
                  -1,
                   Ο,
                   Ο,
                  'BLOSUM62',
                  10,
                   Ο,
                   0,
                   0,
                   0,
                   0)
     order by score, seq id;
```

The output of this table function is shown as follows.

SEQ_ID	SCORE	EVALUE
P58107	49	7.24297332
P31947	169	8.8130E-14
P27348	198	3.8228E-17
Q04917	198	3.8228E-17
P31946	205	5.8977E-18

Index

A	Data Mining SQL functions, 4-13
add-in, 3-3	installation directory in Oracle 10.1, 4-2
APIs, 4-1	installation directory in Oracle 10.2, 4-2
1110, 11	Java, 4-14
D	listing, 4-3
<u>B</u>	obtaining, 4-2
BLAST	PL/SQL, 4-11
data sets, 4-8	set up, 4-1
demo program, 4-20	text mining, 4-18 demo user
table functions, 4-20	creating, 4-7
	granting database privileges, 4-6
C	in Oracle 10.1, 4-5
	in Oracle 10.1, 4-5
CLASSPATH, 4-8	DHCP, 1-4
Custom installation, 1-3, 1-11	directory object, 2-12
	DM\$P_MODEL_EXPIMP_TEMP table, 2-13
D	DM\$P_MODEL_TABKEY_TEMP table, 2-13
Data Miner	DM_NESTED_NUMERICALS, 4-18
see Oracle Data Miner	dmshgrants.sql, 4-6, 4-7
Data Mining option	dmsh.sql, 4-8, 4-18
adding, 1-9	DMSYS, 2-11
installation directory in Oracle 10.1, 1-9	downgrading from release 10.2, 1-17
installation directory in Oracle 10.2, 1-9	dynamic IP, 1-3, 1-4
installing, 1-4	
removing, 1-10	E
Data Mining SQL functions, 4-1, 4-13	<u> </u>
Data Mining user	Enterprise Edition, 1-3
access to data, 2-8	Enterprise Manager, 1-8, 2-1, 2-4
creating in Enterprise Manager, 2-8	environment variables
creating using SQL, 2-7	Linux, 1-15
database privileges, 2-13	Windows, 1-4
Data Pump, 2-12	EXP_FULL_DATABASE privilege, 2-13
Database	expdp command, 2-12
accessing, 1-1	exporting models, 2-11, 4-15
directories, 2-12	
installing, 1-4, 1-5, 1-15	1
privileges, 2-13	IMP ELILI DATADACE
Database Configuration Assistant, 1-3, 1-7, 1-9, 2-3	IMP_FULL_DATABASE privilege, 2-13
Database Control, 1-8, 2-1	impdp command, 2-12
DBA_REGISTRY, 1-14	importing models, 2-11, 4-15
DBMS_DATA_MINING, 2-13, 4-11	initialization files, 1-15
DBMS_DATA_MINING_TRANSFORM, 4-11	installation
demo programs	Advanced, 1-2, 1-11
BLAST, 4-20	Basic, 1-2
data, 4-8	Client, 1-2

Companion, 1-2 Custom, 1-3, 1-11 Database, 1-2 directories, 1-2 directory in 10.1, 1-9 directory in 10.2, 1-9 documentation, 1-2 Enterprise Edition, 1-3 on Linux, 1-15 on Windows, 1-5 Personal Edition, 1-3 preparing for, 1-3 Standard Edition, 1-3 starter database, 1-3, 1-7 verifying, 1-14 iSQL*Plus, 1-8, 2-4	installation directory in Oracle 10.1, 1-9 installation directory in Oracle 10.2, 1-9 installation type, 1-3 installing, 1-4 on Linux, 1-15, 2-3 on Windows, 1-1 removing, 1-10 Scoring Engine, 1-11 SQL functions, 4-13 verifying installation, 1-14 Oracle Data Pump, 2-12 Oracle Enterprise Manager, 2-1 Oracle Net Configuration Assistant, 3-6 Oracle Net Service Name, 3-6 Oracle Services see services
J	Oracle Spreadsheet Add-In for Predictive Analytics, 1-2, 3-3
Java API demo programs, 4-14 javax.datamining.task.ExportTask interface, 2-13 javax.datamining.task.ImportTask utility, 2-13 JDM 1.0 standard, 2-13, 4-14 JSR-73 specification, 2-13, 4-14	Oracle Client, 3-4 Oracle Text, 4-18 Oracle Universal Installer directories, 1-2 installation methods, 1-2 installation types, 1-3
L	platform, 1-2
Linux administration, 2-3 installing Oracle Database on, 1-15 setting environment variables, 1-15	starter database, 1-3 ORACLE_HOME, 1-4, 1-5, 1-9, 2-1 oracle.dmt.jdm.transform.text.OraTextTransform interface, 4-19
log files	Р
export example, 2-14 import example, 2-15	<u>-</u>
loopback adapter, 1-3, 1-4	Personal Edition, 1-3 PL/SQL API demo programs, 4-11 Predictive Analytics, 1-2, 3-3
M	
Microsoft Excel, 3-3	R
N	runInstaller program, 1-15
	S
nested tables, 4-18	-
net service name, 3-6	sample programs see demo programs
0	sample schemas, 1-1, 1-10, 2-3
	Scoring Engine
Oracle Client	adding, 1-13
installing, 3-4 local database, 3-7	installing, 1-11
remote database, 3-6	removing, 1-13
Oracle Data Miner	verifying installation, 1-14 service name, 3-6
connecting to Oracle Database, 3-2	services, 1-4, 1-10, 2-3
installation steps, 3-2	Spreadsheet Add-In for Predictive Analytics, 3-3
installing on Linux, 3-1	SQL functions for data mining, 4-13
installing on WIndows, 3-1	SQL*Plus, 2-2
requirements, 3-2	Standard Edition, 1-3
Oracle Data Mining	starter database, 1-2, 1-7
adding, 1-9	
database administration, 2-1	
demo programs, 4-1	

T

```
tablespaces, 2-4
text mining
demo programs, 4-18
Java, 4-19
PL/SQL, 4-18
```

U

```
Universal Installer

see Oracle Universal Installer

upgrading from release 10.1, 1-15

user

creating for data mining, 2-4, 4-7

see data mining user

see demo user
```

V

V\$OPTION, 1-14

W

```
Windows
administration, 2-1
hardware requirements, 1-3
installing Oracle Database on, 1-1, 1-4
operating system requirements, 1-4
services, 1-4
```