

About the Data Mining API

Overview of the Oracle Data Mining application programming interface (API) components.

- [About Mining Models](#)
- [Data Mining Data Dictionary Views](#)
- [Data Mining PL/SQL Packages](#)
- [Data Mining SQL Scoring Functions](#)

2.1 About Oracle Machine Learning Models

55

Data mining models are database schema objects that perform data mining techniques.

As with all schema objects, access to data mining models is controlled by database privileges. Models can be exported and imported. They support comments and they can be tracked in the Oracle Database auditing system.

Data mining models are created by the `CREATE_MODEL` procedure in the `DBMS_DATA_MINING` PL/SQL package. Models are created for a specific data mining technique, and they use a specific algorithm to perform that function. **Machine learning function** is a term that refers to a class of data mining problems to be solved. Examples of data mining techniques are: regression, classification, attribute importance, clustering, anomaly detection, and feature selection. Oracle Data Mining supports one or more algorithms for each data mining technique.

Along with the data mining technique, in the `CREATE_MODEL` procedure you can specify a settings table to specify an algorithm and other characteristics of a model. Some settings are general, some are specific to a data mining technique, and some are specific to an algorithm.



Note:

Most types of data mining models can be used to score data. However, it is possible to score data without applying a model. Dynamic scoring and predictive analytics return scoring results without a user-supplied model. They create and apply transient models that are not visible to you.

Related Topics

- [Dynamic Scoring](#)
- [DBMS_PREDICTIVE_ANALYTICS](#)
Understand the routines of `DBMS_PREDICTIVE_ANALYTICS` package.
- [Creating a Model](#)
Explains how to create data mining models and query model details.
- [Administrative Tasks for Oracle Data Mining](#)
Explains how to perform administrative tasks related to Oracle Data Mining.

Related Topics

- *Oracle Database PL/SQL Packages and Types Reference*

5.3.5 Specifying Mining Model Settings for R Model

The mining model settings for R model determine the characteristics of the model. You can specify the mining model settings in the `mining_model_table`.

You can build R models with the mining model settings by combining together generic settings that do not require an algorithm, such as `ODMS_PARTITION_COLUMNS` and `ODMS_SAMPLING`. The following settings are exclusive to R mining model, and they allow you to specify the R Mining model:

- `ALGO_EXTENSIBLE_LANG`
- `RALG_BUILD_FUNCTION`
- `RALG_BUILD_PARAMETER`
- `RALG_DETAILS_FORMAT`
- `RALG_DETAILS_FUNCTION`
- `RALG_SCORE_FUNCTION`
- `RALG_WEIGHT_FUNCTION`

Related Topics

- [Registered R Scripts](#)

The `RALG_*_FUNCTION` must specify R scripts that exist in the R script repository. You can register the R scripts using Oracle R Enterprise.

5.3.5.1 ALGO_EXTENSIBLE_LANG

Use the `ALGO_EXTENSIBLE_LANG` setting to specify the Oracle Data Mining framework with extensible algorithms.

Currently, `R` is the only valid value for `ALGO_EXTENSIBLE_LANG`. When the value for `ALGO_EXTENSIBLE_LANG` is set to `R`, the mining models are built using the R language. You can use the following settings in the `model_setting_table` to specify the build, score, and view of the R model.

- `RALG_BUILD_FUNCTION`
- `RALG_BUILD_PARAMETER`
- `RALG_DETAILS_FUNCTION`
- `RALG_DETAILS_FORMAT`
- `RALG_SCORE_FUNCTION`
- `RALG_WEIGHT_FUNCTION`

Related Topics

- [Registered R Scripts](#)

The `RALG_*_FUNCTION` must specify R scripts that exist in the R script repository. You can register the R scripts using Oracle R Enterprise.

1

Introduction to Oracle Data Mining

Introduces Oracle Data Mining to perform a variety of mining tasks.

- [About Oracle Data Mining](#)
- [Data Mining in the Database Kernel](#)
- [Oracle Data Mining with R Extensibility](#)
- [Data Mining in Oracle Exadata](#)
- [About Partitioned Model](#)
- [Interfaces to Oracle Data Mining](#)
- [Overview of Database Analytics](#)

1.1 About Oracle Data Mining

Understand the uses of Oracle Data Mining and learn about different mining techniques.

Oracle Data Mining provides a powerful, state-of-the-art data mining capability within Oracle Database. You can use Oracle Data Mining to build and deploy predictive and descriptive data mining applications, to add intelligent capabilities to existing applications, and to generate predictive queries for data exploration.

Oracle Data Mining offers a comprehensive set of in-database algorithms for performing a variety of mining tasks, such as classification, regression, anomaly detection, feature extraction, clustering, and market basket analysis. The algorithms can work on standard case data, transactional data, star schemas, and text and other forms of unstructured data. Oracle Data Mining is uniquely suited to the mining of very large data sets.

Oracle Data Mining is one of the two components of the **Oracle Advanced Analytics Option** of Oracle Database Enterprise Edition. The other component is Oracle R Enterprise, which integrates R, the open-source statistical environment, with Oracle Database. Together, Oracle Data Mining and Oracle R Enterprise constitute a comprehensive advanced analytics platform for big data analytics.

Related Topics

55

- [Oracle R Enterprise Documentation Library](#)

1.2 Data Mining in the Database Kernel

Learn about implementation of Data Mining.

Oracle Data Mining is implemented in the Oracle Database kernel. Data Mining models are first class database objects. Oracle Data Mining processes use built-in features of Oracle Database to maximize scalability and make efficient use of system resources.

Data mining within Oracle Database offers many advantages:

About Oracle Machine Learning for Python

55

The following topics describe Oracle Machine Learning for Python (OML4Py) and its advantages for the Python user.

- [What Is Oracle Machine Learning for Python?](#)
Oracle Machine Learning for Python (OML4Py) enables you to run Python commands for data transformations and for statistical, machine learning, and graphical analysis on data stored in or accessible through an Oracle database using a Python API. The OML4Py supports running user-defined Python functions through the database spawned and controlled Python engines, with optional built-in data-parallelism and task-parallelism. This embedded execution functionality enables invoking user-defined functions from SQL, and on ADB, REST. The OML4Py supports Automated Machine Learning (AutoML) for algorithm and feature selection, and model tuning and selection. You can augment the Python included functionality with third-party packages from the Python ecosystem.
- [Advantages of Oracle Machine Learning for Python](#)
Using OML4Py to prepare and analyze data in or accessible to an Oracle database has many advantages for a Python user.
- [Manipulate database tables and views using familiar Python functions and syntax](#)
With the transparency layer classes, you can manipulate database tables and views using familiar Python functions and syntax. For example, using DataFrame proxy objects that map to database data, users can invoke overloaded Pandas functions that transparently generate SQL that runs in the database, using the database as a high-performance compute engine.
- [About the Python Components and Libraries in OML4Py](#)
OML4Py requires an installation of Python, the specified Python libraries, as well as the OML4Py components.

2.1 What Is Oracle Machine Learning for Python?

Oracle Machine Learning for Python (OML4Py) enables you to run Python commands for data transformations and for statistical, machine learning, and graphical analysis on data stored in or accessible through an Oracle database using a Python API. The OML4Py supports running user-defined Python functions through the database spawned and controlled Python engines, with optional built-in data-parallelism and task-parallelism. This embedded execution functionality enables invoking user-defined functions from SQL, and on ADB, REST. The OML4Py supports Automated Machine Learning (AutoML) for algorithm and feature selection, and model tuning and selection. You can augment the Python included functionality with third-party packages from the Python ecosystem.

OML4Py is a Python module that enables Python users to manipulate data in database tables and views using Python syntax. OML4Py functions and methods transparently translate a select set of Python functions into SQL for in-database execution.

OML4Py is available in the following Oracle database environments:

- OML4Py is available in the Python interpreter in Oracle Machine Learning Notebooks in your Oracle Autonomous Database. For more information, see *Get Started with Notebooks for Data Analysis and Data Visualization* in *Using Oracle Machine Learning Notebooks*.

Install OML4Py for On-Premises Databases

The following topics tell how to install and uninstall the server and client components required for using OML4Py with an on-premises Oracle Database.

Topics:

- [OML4Py On Premises System Requirements](#)
OML4Py on premises runs on 64-bit platforms only.
- [Build and Install Python for Linux for On-Premises Databases](#)
Instructions for installing Python for Linux for an on-premises Oracle database.
- [Install the Required Supporting Packages for Linux for On-Premises Databases](#)
Both the OML4Py server and client installations for an on-premises Oracle database require that you also install a set of supporting Python packages, as described below.
- [Install OML4Py Server for On-Premises Oracle Database](#)
The following instructions tell how to install and uninstall the OML4Py server components for an on-premises Oracle Database 23ai.
- [Install OML4Py Client for On-Premises Oracle Database](#)
Instructions for installing and uninstalling the on-premises OML4Py client.

4.1 OML4Py On Premises System Requirements

OML4Py on premises runs on 64-bit platforms only.

Both client and server on-premises components are supported on the Linux platforms listed in the table below.

Table 4-1 On-Premises OML4Py Platform Requirements

| Operating System | Hardware Platform | Description |
|-------------------------|-------------------|-------------------------------|
| Oracle Linux x86-64 8.x | Intel | 64-bit Oracle Linux Release 8 |

Table 4-2 On-Premises OML4Py Configuration Requirements and Server Support Matrix

| Oracle Machine Learning for Python Version | Python Version | On-Premises Oracle Database Release |
|--|----------------|-------------------------------------|
| 2.0 | 3.12.0 | 19c, 21c, 23ai |

55



Note:

Plug in violations are seen after DB upgrade from 21c to 23c with OML4PY configurations.