

# ORACLE DATABASE 11G

## ORACLE DATA MINING オプション

### 主な機能と強化

**ORACLE** 11g  
DATABASE

#### Oracle Data Mining

- In-Database アナリティクス
- データ・マイニング・モデル (最高級のデータベース・オブジェクト)
- アルゴリズムの完備
- 異常検出
- 自動データ準備
- マイニング・アクティビティ・ガイド
- データ準備
- テキスト・マイニング
- Java API および PL/SQL API
- コード生成
- 予測解析
- Oracle Database プラットフォーム

Oracle Database 11g Release 1 の Data Mining オプションを使用すると、企業は、対応可能な予測情報を生成し、統合されたビジネス・インテリジェンス・アプリケーションを構築できます。ビジネス・アナリストは、データに隠されたパターンや見識を発見できます。アプリケーション開発者は、新しいビジネス・インテリジェンス (予測、パターン、および発見) の発見と企業全体への展開を迅速に自動化できます。

#### In-Database アナリティクス

Oracle Data Mining (ODM) は、Oracle Database に埋め込まれた総合的なデータ・マイニング機能を提供します。そのため、データを分析するためにデータベースから外部の分析エンジンにデータを抽出する必要がありません。Oracle Database 11g は、Oracle Data Mining のすべての機能を搭載しています。Oracle Data Mining では、データはデータベースの外部へ出ることがありません。データ、データ準備、モデル構築、およびモデル・スコアリング処理のすべては Oracle Database 内で行われます。そのため、セキュリティ、拡張性、管理性、アプリケーション開発、およびユーザー・アクセスに関しても、大きな利点があります。

Oracle Data Mining のデータベースに組み込まれたデータ・マイニング機能により、データがデータベースに残るだけでなく、マイニング・タスクやデータ変換もデータベース内部で実行されます。任意のユーザー・インタフェースから独立して、自動的、非同期的に実行できます。

Oracle Database 11g の拡張性により、Oracle Data Mining は、大量のデータを分析して曖昧なパターンや関係を検出し、データに隠された新しいビジネス・インテリジェンスを抽出できます。新しい見識や予測はデータベースの表に格納され、他の Oracle や Oracle 以外からの問合せ、分析、レポート用のツールやアプリケーションによりアクセスできます。

#### アルゴリズムの完備

Oracle Data Mining は、分類、回帰、クラスタリング、相関分析、異常検出、テキスト・マイニング、属性評価、特徴抽出のような、さまざまなデータ・マイニングのモデル構築および評価機能をサポートします。

分類問題を解決するために、ディシジョン・ツリー、Naïve Bayes、一般化線形回帰 (GLM)、およびサポート・ベクター・マシン (SVM) のようなアルゴリズムがサポートされます。

拡張された K-means（距離メトリックに依存）または O-Cluster（密度に依存）を使用すると、クラスタリングを実行できます。

一般化線形回帰またはサポート・ベクター・マシンを使用すると、回帰問題を解決できます。

テキスト・マイニングと異常検出はサポート・ベクター・マシンを使用し、属性評価は最小記述長（MDL）を使用し、相関分析は A Priori を使用し、特徴抽出は Non-Negative Matrix Factorization（NMF）を使用します。

### 異常検出

通常、分類では、ターゲットクラスの知識が必要です。SVM のバージョンは 1 つのクラスのプロファイルを構築できます。これが適用されると、そのプロファイルと異なる（すなわち"異常"または"不審"）条件がフラグで指摘されます。そのため、分類を使用してもほとんど特定できないような、必ずしも相互に関係がないまれな条件を検出できます。

### 自動データ準備

データ・マイニング・モデルに必要なデータ変換処理は、通常、手動で実行するか、マイニング・アクティビティの一部として実行する必要があります。Oracle 11g Data Mining では、特定のアルゴリズムに必要なすべてのデータ変換が、モデル実行の一部として自動的に処理されます。

### マイニング・アクティビティ

Oracle Data Miner のグラフィカル・ユーザー・インタフェース（GUI）は、操作の正しい順序を設定し、アルゴリズムに必要なすべてのデータ変換を実行するだけでなく、すべてのパラメータに対して優れた設定と最適化を提供するマイニング・アクティビティを採用しています。ただし、熟練ユーザーに対しては、デフォルト値を上書きするためにすべてのパラメータを公開できます。

モデル構築作業には、モデル評価（適宜）、および分類用の受信者操作特性（ROC）分析や回帰用の残差プロットのような複数のテスト方法が含まれます。

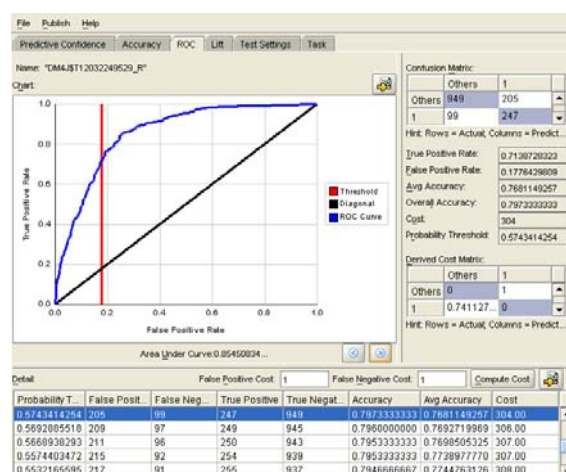


図 1 受信者操作特性（ROC）分析

モデル構築作業は、すべてのデータ変換とパラメータ設定を"記憶"するので、最適モデルによるデータのスコアリング時には、自動実行のために構築時のメタデータがモデル適用作業にシームレスに渡されます。

### データ準備

Oracle Data Miner は、入力として複数の表やビューを受け入れ、モデリングに必要な適切な結合や変換を実行できます。トランザクション・データやネストされたデータ表をマイニングすることができます。データベースの内部でデータ集計、変換およびデータ準備処理が管理されるので、モデルの配置やアプリケーションの開発が高速化されます。

### テキスト・マイニング

サポート・ベクター・マシン、相関分析、K-means のクラスタリング、および Non-Negative Matrix Factorization アルゴリズムのすべてが、入力属性としてテキスト（構造化されていないデータ）を受け入れるため、たとえば、医師のメモ、テクニカル・ペーパー、または警察報告書などを含む列を他の任意の入力変数と同様に処理できるので、予測モデルの価値が上昇します。

### Java API または PL/SQL API

アプリケーションの開発者は、Oracle Data Mining の Java または PL/SQL アプリケーション・プログラミング・インタフェース（API）を使用して、データ・マイニングの見識と予測をビジネス・アプリケーションに統合できます。開発者は、いずれかの API（たとえば PL/SQL API）を使用して、データ・マイニング・モデルを構築し、そのモデルを適用できます。Oracle Database には、ほとんどのデータ・マイニング操作に必要なコーディングを示すサンプル・プログラムが付属しています。

### コード生成

Oracle Data Miner の GUI が作業の操作を実行すると、PL/SQL コードが生成されるので、予測モデルを生成し、テストし、または適用して別の Oracle Database 環境にパッケージ化し、実行できます。コードは、Oracle JDeveloper や Oracle SQLDeveloper からアクセスでき、アプリケーションの作成に使用できます。そのため、1 つのシステムで構築され、最適化されたモデルを、別のシステムでアプリケーションのコンポーネントとしてデータに適用できます。

### 予測解析

SQL の PREDICT 関数、EXPLAIN 関数、および PROFILE 関数は、分類モデルまたは属性評価モデルを構築するための自己完結型パッケージです。すべてのパラメータに最適化された値が入力され、中間データは保存されません。その結果は、予測されたスコア（PREDICT）、属性の順リスト（EXPLAIN）、またはスコアとルール（PROFILE）であり、演算パイプラインの一部として使用したり、コマンド・ラインやスプレッドシートに表示したりできます。

	A	B	C	E	F	G	H
1	AGE	WORKCLASS	ANNUAL_INCOME	S_SINCE_LAST_PURCH	MARITAL_STATUS	OCCUPATION	HOUSEHOLD_SIZE
2	41	Private	19645		14 NeverM	Prof.	2
3	27	Private	177351		13 NeverM	Sales	2
4	20	Private	154781		9 NeverM	Cleric.	2
5	45	SelfEI	34091		13 Married	Exec.	3
6	34	Private	265706		14 NeverM	Sales	9+
7	38	Private	255621		9 Married	Other	3
8	28	Private	206351		10 Married	Sales	3
9	19	Private	259352		9 NeverM	Sales	2

図 2 スプレッドシートでの PREDICT 関数の実行

### Oracle Database プラットフォーム

Oracle Data Mining を使用すると、企業は、完全に統合された Oracle Data Warehouse、Oracle Business Intelligence 環境を利用できます。Oracle Data Mining のすべての機能は、業界をリードする Oracle Database のセキュリティ、拡張性およびデータ管理プラットフォームに統合されています。