

ORACLE® Oracle Databaseでの機械学習 - 何をしたいですか?

Oracle Machine Learningにより、AIアプリケーションとダッシュボードを構築し、強力なインデータベースMLアルゴリズム、自動ML機能、オープン・ソースのPythonおよびRとの統合を実現できます。Oracle Machine Learningのアルゴリズムは、パラレル実行に対応することによってパフォーマンスとスケーラビリティを高め、メモリ使用率を改善し、パーティション化モデルとテキスト列の自動マイニングに対応できます

分類

2つ（バイナリ）以上（マルチクラス）のカテゴリ値を含むターゲット変数を予測

ディジション・ツリー	人が解釈可能なルールを生成、セグメント化で使用可能
明示的セマンティック分析	大量テキスト・コーパスに適したテキスト・カテゴリ化
極端な勾配ブーステッド	よく使用されるXGBoostアルゴリズムのスケラブル実装、ツリー・モデルと線形モデルをサポート
ロジスティック回帰/ 一般化線形モデル	属性係数とモデル統計によるバイナリ（0/1、Yes/No）ターゲット属性の予測、ナロー/ワイド/スパース・データ、リッジおよび特徴選択/生成の有効化、行診断
ナイーブ・ベイズ	条件付き確率を計算し、解析可能な確率を算出、予測子属性の独立を仮定
ニューラル・ネットワーク	ノイズの多い複雑なデータに最適、多くの隠れレイヤーに対応
ランダム・フォレスト	バギングと特徴ランダムネスに依存するツリーベース・アンサンブル法
サポート・ベクター・マシン	線形および非線形問題の解決、マルチ・ソルバー、疎最適化、マルチターゲット分類のサポート（行あたりのターゲット・リスト）

回帰

数値ターゲット変数を予測

極端な勾配ブーステッド	よく使用されるXGBoostアルゴリズムのスケラブル実装、ツリー・モデルと線形モデルをサポート
一般化線形モデル	属性により数値ターゲット属性を予測 係数とモデル統計、ナロー/ワイド/スパース・データ、リッジな特徴選択/生成の有効化、行診断
ニューラル・ネットワーク	ノイズの多い複雑なデータに最適、多くの隠れレイヤーに対応
ステップワイズ回帰	線形モデルで予測子の"最適"セットを選択、前進、後退、両方、代替方向をサポート
サポート・ベクター・マシン	線形および非線形問題の解決、マルチ・ソルバー、疎最適化

属性の重要度

教師付きおよび教師なしの変数ランキングによるモデル品質の向上

CUR分解	教師なし手法として、属性の重要度をランキングするための低ランクSVDベースのアプローチをサポート
期待値最大化	教師なし変数ランキングとペアワイズ依存性の推定をサポート
最小記述長	分類と回帰にもっとも重要な変数を選択

ランキング

他のアイテムに対する1つのアイテムのランキングの教師付き予測可能性

極端な勾配ブーステッド	ペアワイズとリストワイズのランキングをサポート
-------------	-------------------------

クラスタ化

ケースを階層型クラスタにグループ化またはセグメント化して、確率、ルール、統計を生成

期待値最大化	自動化モデル検索、過剰適合からの保護、数値と多項分布、高性能確率推定
k-means	クラスタの指定数kを算出 ユークリッド関数と余弦距離関数、疎最適化
直交パーティション化	指定された最大数まで自然クラスタを検出、密度ベース

特徴抽出

すべての入力変数によって少ない数の変数セットが生成されると見なされる新しい値を導出

明示的セマンティック分析	コーパスから派生した、人が読取り可能なトピック・ラベルによるテキストのカテゴリ化、ドキュメント間でセマンティック類似度を推定
非負マトリックス因数分解	非負線形結合に基づく特徴の導出により特徴の解析可能性を向上
主成分分析	SVDを使用して、データセットからの最大変数を含む非相関変数のセットを取得
特異値分解	ツールおよびスキニー・ソルバーを介するナロー・データ、確率的ソルバーを介するワイド・データ

異常検出

正常なデータのパターンを学習することにより、正常または異常としてケースを識別

One-Class SVM	ターゲットを使用しないSVM分類の特殊ケース、線形および非線形問題の解決、複数ソルバー、疎最適化
MSET-SPRT	IoTセンサー・データで、非線形の非パラメータ・パターンにより異常を検出するためのプロセス監視、"多変量状態推定手法"

時系列

数値型または日付/タイムスタンプ型のいずれかの連続オーダー列を使用して順次数値データを予想または予測

指数平滑法	トレンドと季節性のある場合とない場合の規則的系と不規則系向けのシングル、ダブル、トリプル指数平滑法、Holt-Wintersを含む複数の手法に対応
-------	---------------------------------------------------------------------------

関連付けルール

トランザクションまたは2Dデータ表現を使用した市場基本分析により、頻繁に発生するパターンとルールを抽出

Apriori	頻出アイテムセットを検出して人が解釈可能なルールを生成、ルールに関連付けられたサポート、確信、リフト、集約の各手法を計算
---------	--------------------------------------------------------------

行の重要性

行の教師なしランキング

CUR分解	教師なし手法として、行の重要度をランキングするための低ランクSVDベースのアプローチをサポート
-------	-------------------------------------------------