Oracle Machine Learning Notebook
Included in Autonomous Data Warehouse Cloud

Charlie Berger, MS Engineering, MBA
Sr. Director Product Management,
Machine Learning, AI and Cognitive Analytics
charlie.berger@oracle.com
www.twitter.com/CharlieDataMine
Safe Harbor Statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle’s products remains at the sole discretion of Oracle.
Introducing Oracle Autonomous Data Warehouse Cloud

Value Proposition

**Easy**
- Provision a data warehouse in as little as 15-seconds
- **Automated management** of database administration
- Simple Load and Go with **Automated Tuning**
- Dedicated cloud-ready **migration tools including Redshift**

**Fast**
- **Up to 14x performance advantage** than Redshift\(^1\)
- High **concurrency** supports multi-user access and workloads
- Based on **Exadata** for extreme performance

**Only Pay for What you Use** with user defined sizing, on-demand scaling & idle shut-off
- **Independent scaling** of compute and storage
- Instant scaling with **zero downtime**
Oracle Autonomous Data Warehouse Cloud Key Features

High-Performance Queries and Concurrent Workloads
Optimized query performance with preconfigured resource profiles for different types of users

Oracle SQL
Autonomous DW Cloud is compatible with all business analytics tools that support Oracle Database

Self Driving
Fully automated database for self-tuning patching and upgrading itself while the system is running

Cloud-Based Data Loading
Fast, scalable data-loading from Oracle Object Store, AWS S3, or on-premises

Highly Elastic
Independently scale compute and storage, without having to overpay for fixed blocks of resources

Built-in Web-Based SQL ML Tool
Apache Zeppelin Oracle Machine Learning notebooks ready to run ML from browser

Database migration utility
Dedicated cloud-ready migration tools for easy migration from Amazon Redshift, SQL Server and other databases

Enterprise Grade Security
Data is encrypted by default in the cloud, as well as in transit and at rest

Oracle Machine Learning
Architecture for Modern Cloud Data Warehousing

Autonomous Data Warehouse Cloud
- Service Management
- Built-in Access Tools
- Oracle Machine Learning
- Oracle Object Storage Cloud
- Flat Files and Staging

Autonomous Database Cloud

Developer Tools
- Oracle SQL Developer

Data Integration Services
- Oracle Data Integration Platform Cloud
- 3rd Party DI on Oracle Cloud Compute
- 3rd Party DI On-premises

Advanced Analytics
- Oracle Analytics Cloud
- 3rd Party Analytics on Oracle Cloud Compute
- 3rd Party Analytics On-premises

Oracle Database Cloud Services
- EDWs, DW, departmental marts and sandboxes

Copyright © 2018, Oracle and/or its affiliates. All rights reserved.
Introducing: Oracle Machine Learning SQL Notebook
Oracle Machine Learning
Machine Learning Notebook for Autonomous Data Warehouse Cloud

Key Features

• Collaborative UI for data scientists
  – Packaged with Autonomous Data Warehouse Cloud (V1)
  – Easy access to shared notebooks, templates, permissions, scheduler, etc.
  – SQL ML algorithms API (V1)
  – Supports deployment of ML analytics
Oracle Machine Learning

Machine Learning Notebook for Autonomous Data Warehouse Cloud

Key Features

• Collaborative UI for data scientists
  – Packaged with Autonomous Data Warehouse Cloud (V1)
  – Easy access to shared notebooks, templates, permissions, scheduler, etc.
  – SQL ML algorithms API (V1)
  – Supports deployment of ML analytics
Oracle Machine Learning Quick DEMO
Sign In

- Tenant: TENANT1
- Database: PDB1
- Username: CBERGER
- Password: **********

Sign In
<table>
<thead>
<tr>
<th>Name</th>
<th>Owner</th>
<th>Type</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Charlie Workspace</td>
<td>CBERGER</td>
<td>Workspace</td>
<td></td>
</tr>
<tr>
<td>Charlie Project</td>
<td>CBERGER</td>
<td>Project</td>
<td></td>
</tr>
<tr>
<td>Customer Analytics</td>
<td>CBERGER</td>
<td>Project</td>
<td>Using Machine Learning to target top Aff...</td>
</tr>
</tbody>
</table>
Create Project

Name
Customer Analytics

Comment
Using Machine Learning to target top Affinity Card responders

Select Workspace:
Charlie Workspace / CBERGER

OK Cancel
<table>
<thead>
<tr>
<th>Name</th>
<th>Last Update</th>
<th>Updated By</th>
<th>Connection Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anomaly Detection</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Association Rules</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Attribute Importance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classification Prediction Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clustering</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My First Notebook</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Regression _1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQL Query Scratchpad</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQL Script Scratchpad</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SQL Statistical Functions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2/8/18 1:00 PM</td>
<td>CBERGER</td>
<td>Global</td>
</tr>
<tr>
<td></td>
<td>2/8/18 1:00 PM</td>
<td>CBERGER</td>
<td>Global</td>
</tr>
<tr>
<td></td>
<td>2/8/18 1:00 PM</td>
<td>CBERGER</td>
<td>Global</td>
</tr>
<tr>
<td></td>
<td>2/8/18 1:00 PM</td>
<td>CBERGER</td>
<td>Global</td>
</tr>
<tr>
<td></td>
<td>2/8/18 1:00 PM</td>
<td>CBERGER</td>
<td>Global</td>
</tr>
<tr>
<td></td>
<td>2/8/18 1:00 PM</td>
<td>CBERGER</td>
<td>Global</td>
</tr>
<tr>
<td></td>
<td>2/8/18 1:00 PM</td>
<td>CBERGER</td>
<td>Global</td>
</tr>
<tr>
<td>CUST_ID</td>
<td>CUST_FIRST_NAME</td>
<td>CUST_LAST_NAME</td>
<td>CUST_GENDER</td>
</tr>
<tr>
<td>---------</td>
<td>----------------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>43850</td>
<td>Rosemary</td>
<td>Kane</td>
<td>Male</td>
</tr>
<tr>
<td>52722</td>
<td>Rosemary</td>
<td>Kane</td>
<td>Male</td>
</tr>
<tr>
<td>46784</td>
<td>Rosemary</td>
<td>Kane</td>
<td>Male</td>
</tr>
<tr>
<td>3895</td>
<td>Rosemary</td>
<td>Kane</td>
<td>Male</td>
</tr>
<tr>
<td>4750</td>
<td>Rosemary</td>
<td>Kane</td>
<td>Male</td>
</tr>
<tr>
<td>14562</td>
<td>Rosemary</td>
<td>Kane</td>
<td>Male</td>
</tr>
</tbody>
</table>

SQL Script Scratchpad

```
SELECT * FROM SH_CUSTOMERS;
```
<table>
<thead>
<tr>
<th>CUST_ID</th>
<th>CUST_FIRST_NAME</th>
<th>CUST_LAST_NAME</th>
<th>CUST_GENDER</th>
<th>CUST_YEAR_OF_BIRTH</th>
<th>CUST_MARITAL_STATUS</th>
<th>CUST_STREET_ADDRESS</th>
<th>CUST_POSTAL_CODE</th>
<th>CUST_CITY</th>
<th>CUST_CITY_ID</th>
<th>CUST_STATE_PROVINCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>37.057</td>
<td>Bernard</td>
<td>Wright</td>
<td>M</td>
<td>1,941</td>
<td>married</td>
<td>107 East Catano Avenue</td>
<td>66,361</td>
<td>Vebert</td>
<td>52,436</td>
<td>Northern-Westfalen</td>
</tr>
<tr>
<td>40.612</td>
<td>Bernard</td>
<td>Wright</td>
<td>M</td>
<td>1,947</td>
<td></td>
<td>107 South Prentiss Avenue</td>
<td>33,866</td>
<td>Bergen op Zoom</td>
<td>51,181</td>
<td>Noord-Brabant</td>
</tr>
<tr>
<td>44.167</td>
<td>Bernard</td>
<td>Wright</td>
<td>M</td>
<td>1,974</td>
<td>single</td>
<td>117 North Door Avenue</td>
<td>83,601</td>
<td>San Francisco</td>
<td>52,289</td>
<td>CA</td>
</tr>
<tr>
<td>34.880</td>
<td>Bernard</td>
<td>Wright</td>
<td>M</td>
<td>1,955</td>
<td>married</td>
<td>17 North Lehigh Court</td>
<td>59,862</td>
<td>Malaga</td>
<td>51,864</td>
<td>Malaga</td>
</tr>
<tr>
<td>47.945</td>
<td>Bernard</td>
<td>Wright</td>
<td>M</td>
<td>1,937</td>
<td>married</td>
<td>27 West Baraga Boulevard</td>
<td>46,864</td>
<td>Lauret</td>
<td>51,768</td>
<td>Languedoc-Roussillon</td>
</tr>
<tr>
<td>1,502</td>
<td>Bernard</td>
<td>Wright</td>
<td>M</td>
<td>1,945</td>
<td>single</td>
<td>37 Mountain View Street</td>
<td>80,841</td>
<td>Wolverhampton</td>
<td>52,514</td>
<td>England - West Midlands</td>
</tr>
<tr>
<td>5,057</td>
<td>Bernard</td>
<td>Wright</td>
<td>M</td>
<td>1,947</td>
<td></td>
<td>37 South Catano Street</td>
<td>34,216</td>
<td>Munich</td>
<td>51,934</td>
<td>Bayern</td>
</tr>
<tr>
<td>8,612</td>
<td>Bernard</td>
<td>Wright</td>
<td>M</td>
<td>1,939</td>
<td></td>
<td>47 West Prentiss Road</td>
<td>72,059</td>
<td>Los Angeles</td>
<td>51,806</td>
<td>CA</td>
</tr>
</tbody>
</table>
Simple Oracle Machine Learning notebook example

Oracle Machine Learning example notebook for learning basic functions using SH schema data and highlights basic data selection and data viewing using the Oracle Autonomous Data Warehouse Cloud (ADWC).

By Charlie Berger


Show all tables

| TABLE_NAME | CLUSTER_NAME | IOT_NAME | STATUS | PCT_FREE | PCT_USED | INIT_TRANS | MAX_TRANS | INITIAL_EXTENT | NEXT_EXTENT | MIN_EXTENTS | MAX_EXTENTS | PCT_INCREASE | FREELISTS | FREELIST_GROUPS | LOGGING | BACKED_UP |
|------------|--------------|----------|--------|----------|----------|-------------|------------|----------------|-------------|-------------|--------------|--------------|-------------|------------|-----------------|---------|-----------|
| SALES      | SYSTEM       |          | VALID  | 10       | 90       | 1           | 255        | 1              | 255         | 1           | 1            | 1            | YES         | NO          |                 |         |           |
| TIMES      | SYSTEM       |          | VALID  | 10       | 90       | 1           | 255        | 1              | 255         | 1           | 1            | 1            | YES         | NO          |                 |         |           |
| CHANNELS   | SYSTEM       |          | VALID  | 10       | 90       | 1           | 255        | 1              | 255         | 1           | 1            | 1            | YES         | NO          |                 |         |           |
| PROMOTIONS | SYSTEM       |          | VALID  | 10       | 90       | 1           | 255        | 1              | 255         | 1           | 1            | 1            | YES         | NO          |                 |         |           |
| CUSTOMERS  | SYSTEM       |          | VALID  | 10       | 90       | 1           | 255        | 1              | 255         | 1           | 1            | 1            | YES         | NO          |                 |         |           |
| COUNTRIES  | SYSTEM       |          | VALID  | 10       | 90       | 1           | 255        | 1              | 255         | 1           | 1            | 1            | YES         | NO          |                 |         |           |
| SUPPLEMENTARY_DEMOGRAPHICS | SYSTEM |          | VALID  | 10       | 90       | 1           | 255        | 1              | 255         | 1           | 1            | 1            | YES         | NO          |                 |         |           |
| SALES_TRANSACTIONS | EXT | SYSTEM | VALID  | 10       | 90       | 1           | 255        | 1              | 255         | 1           | 1            | 1            | YES         | NO          |                 |         |           |
Anomaly Detection to Detect Suspicious or Rare Occurrences

This notebook shows how to detect rare records, customers or transactions using an unsupervised learning algorithm (1-Class Support Vector Machine). The notebook first builds a 1-Class SVM model and then applies the model to flag unusual or suspicious records. The anomaly detection model can also be applied to “score” new records. The entire machine learning methodology runs inside the Oracle Autonomous Data Warehouse Cloud (ADWC).

By Charlie Berger


Clean up and drop any table if previously exists for notebook reproducibility

```sql
BEGIN
EXECUTE IMMEDIATE 'DROP Table SUPPLEMENTARY_DEMOGRAPHICS2';
EXCEPTION
WHEN OTHERS THEN NULL;
END;

PL/SQL procedure successfully completed.
```

Create SUPPLEMENTARY_DEMOGRAPHICS2 table that remove COMMENTS unstructured data for simplicity.

```sql
CREATE Table SUPPLEMENTARY_DEMOGRAPHICS2
AS
(SELECT AFFINITY_CARD, BOOKEEPING_APPLICATION, BULK_PACK_DISKETTES, CUST_ID, EDUCATION, FLAT_PANEL_MONITOR, HOME_THEATER_PACKAGE, HOUSEHOLD_SIZE, OCCUPATION, OS_DOC_SET_KANJI, PRINTER SUPPLIES, YRS_RESIDENCE, Y_BOX_GAMES
FROM SH.SUPPLEMENTARY_DEMOGRAPHICS);

Updated 4500 row(s).
```
Anomaly Detection

Build Anomaly Detection Model (1-Class SVM)

```sql
-- Build Anomaly Detection Model (1-Class SVM) on SUPPLEMENTARY_DEMOGRAPHICS2 data

DECLARE
  v_sql varchar2(100);
BEGIN
  -- drop build settings
  BEGIN
    v_sql := 'DROP TABLE CUSTOMERS360_SET';
    EXECUTE IMMEDIATE v_sql;
    DBMS_OUTPUT.PUT_LINE (v_sql || ' succeeded');
  EXCEPTION WHEN OTHERS THEN
    DBMS_OUTPUT.PUT_LINE ('drop unnecessary - no table exists');
  END;
  -- drop any previous model.
  BEGIN
    v_sql := 'CALL DBMS_DATA_MINING.DROP_MODEL(''CUSTOMERS360MODEL'')';
    EXECUTE IMMEDIATE v_sql;
    DBMS_OUTPUT.PUT_LINE (v_sql || ' succeeded');
  EXCEPTION WHEN OTHERS THEN
    DBMS_OUTPUT.PUT_LINE ('drop unnecessary - no model exists');
  END;
  -- Create a Build Setting table for Model Build
  EXECUTE IMMEDIATE 'CREATE TABLE CUSTOMERS360_SET (setting_name VARCHAR2(30),setting_value VARCHAR2(4000));
  EXECUTE IMMEDIATE 'INSERT INTO CUSTOMERS360_SET (setting_name, setting_value) VALUES (''ALGO_NAME'', ''ALGO_SUPPORT_VECTOR_MACHINES'');
  EXECUTE IMMEDIATE 'INSERT INTO ml_build_settings (setting_name, setting_value) VALUES (''PREP_AUTO'', '''ON''');
  DBMS_OUTPUT.PUT_LINE ('Created model build settings table: CUSTOMERS360_SET');
  -- Build the 1-Class SVM model.
  EXECUTE IMMEDIATE CALL DBMS_DATA_MINING.CREATE_MODEL(''CUSTOMERS360MODEL'', '''CLASSIFICATION'', '''CUSTOMERS360'', '''CUST_ID'' null, '''CUSTOMERS360_SET'');
  DBMS_OUTPUT.PUT_LINE ('Created model: CUSTOMERS360_MODEL');
END;
```

DROP TABLE CUSTOMERS360SET: succeeded
CALL DBMS_DATA_MINING.DROP_MODEL(''CUSTOMERS360MODEL''): succeeded
Created model build settings table: CUSTOMERS360_SET
Created model: CUSTOMERS360_MODEL
PL/SQL procedure successfully completed.
Anomaly Detection

Display CUSTOMERS360 table

SQL

```sql
-- CUST_YEAR_OF_BIRTH vs. YRS_RESIDENCE grouped by CUST_MARITAL_STATUS
SELECT * from CUSTOMERS360;
```

Build Anomaly Detection Model (1-Class SVM)

```sql
-- Build Anomaly Detection Model (1-Class SVM) on SUPPLEMENTARY_DEMOGRAPHICS.csv data
DECLARE v_sql varchar(1000);
BEGIN
  -- drop build settings
  BEGIN
    v_sql := 'DROP TABLE CUSTOMERS_S360_SET';
    EXECUTE IMMEDIATE v_sql;
    DBMS_OUTPUT.PUT_LINE (v_sql || ' succeeded');
  EXCEPTION
    WHEN OTHERS THEN
      DBMS_OUTPUT.PUT_LINE (v_sql || ' drop unnecessary - no table exists');
  END;
```

---

DECLARATION OF sql:

```sql
DECLARE v_sql varchar(1000);
```

BEGIN:

- Drop table or database settings

```sql
v_sql := 'DROP TABLE CUSTOMERS_S360_SET';
EXECUTE IMMEDIATE v_sql;
DBMS_OUTPUT.PUT_LINE (v_sql || ' succeeded');
```

- Exception handling

```sql
EXCEPTION
  WHEN OTHERS THEN
    DBMS_OUTPUT.PUT_LINE (v_sql || ' drop unnecessary - no table exists');
```

---

The diagram shows a scatter plot with data points grouped by marital status, indicating anomalies in the CUSTOMERS360 table.
Anomaly Detection

Graph Customers and probability of being Anomalous

Display the Top 5 Most Anomalous Customers

```
select * from
(select CUST_ID, HOUSEHOLD_SIZE, YRS_RESIDENCE, CUST_GENDER, CUST_MARITAL_STATUS, round(prob_fraud*100,1) percent_fraud,
rank() over (order by prob_fraud desc) rnk from
(select CUST_ID, HOUSEHOLD_SIZE, YRS_RESIDENCE, CUST_GENDER, CUST_MARITAL_STATUS, prediction_probability(CUSTOMERS365MODEL, '0' using *) prob_fraud
from CUSTOMERS50000
)
where rnk <= 5
order by percent_fraud desc
```

<table>
<thead>
<tr>
<th>CUST_ID</th>
<th>HOUSEHOLD_SIZE</th>
<th>YRS_RESIDENCE</th>
<th>CUST_GENDER</th>
<th>CUST_MARITAL_STATUS</th>
<th>PERCENT_FRAUD</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.1099</td>
<td>2</td>
<td>2</td>
<td>F</td>
<td>Widowed</td>
<td>75.73</td>
</tr>
<tr>
<td>103.154</td>
<td>2</td>
<td>2</td>
<td>F</td>
<td>Widowed</td>
<td>75.73</td>
</tr>
<tr>
<td>102.048</td>
<td>9+</td>
<td>2</td>
<td>F</td>
<td>Widowed</td>
<td>73.32</td>
</tr>
<tr>
<td>101.127</td>
<td>9+</td>
<td>3</td>
<td>F</td>
<td>Widowed</td>
<td>66.66</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>75.73</td>
</tr>
</tbody>
</table>
Classification Prediction Model

Predicting Target Customers using Classification

Example notebook to predict customers most likely to be positive responders to an Affinity Card loyalty program. This notebook builds and applies classification models (decision tree) using the SH schema data and processed inside the Oracle Autonomous Data Warehouse Cloud (ADWC).

By Charlie Berger


Display the SH.SUPPLEMENTARY_DEMOGRAPHICS data

```
<table>
<thead>
<tr>
<th>CUST_ID</th>
<th>EDUCATION</th>
<th>OCCUPATION</th>
<th>HOUSEHOLD_SIZE</th>
<th>YRS_RESIDENCE</th>
<th>AFFINITY_CARD</th>
<th>BULK_PACK_DISKETTES</th>
<th>FLAT_PANEL_MONITOR</th>
<th>HOME_THEATER_PACKAGE</th>
<th>BOOKKEEPING_APPLICATION</th>
<th>PRINTER_SUPPLIES</th>
<th>Y_BOX_GAMES</th>
<th>OS_DOCUMENTATION</th>
<th>KANJII_CODE</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.001</td>
<td>&lt; Bach</td>
<td>Crafts</td>
<td>3</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Sh</td>
</tr>
<tr>
<td>100.002</td>
<td>HS-grad</td>
<td>Machine</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Aff</td>
</tr>
<tr>
<td>100.003</td>
<td>HS-grad</td>
<td>Sales</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>Aff</td>
</tr>
</tbody>
</table>
```
Classification Prediction Model

Real-time prediction

SQL:
```sql
SELECT prediction_probability(Material_CLASS_MODEL, '1',
   USING '3' as HOUSEHOLD_SIZE, '1' as YRS_RESIDENCE, '1' as Y_BOX_GAMES)
FROM dual
```

PREDICTION_PROBABILITY(Material_CLASS_MODEL,'1' Using '3' AS HOUSEHOLD_SIZE, '1' AS YRS_RESIDENCE, '1' AS Y BOX_GAMES)

0.450058

Interactive selection of likely Affinity_Card responders selected by HOUSEHOLD_SIZE

SQL:
```sql
SELECT A.*, B.* FROM ML APPLY_RESULT A, ML_TEST_DATA B WHERE HOUSEHOLD_SIZE = $[HOUSEHOLD_SIZE='1','1' OR '1'] and A.CUST_ID = B.CUST_ID;
```

<table>
<thead>
<tr>
<th>CUST_ID</th>
<th>PREDICTION</th>
<th>PROBABILITY</th>
<th>COST</th>
<th>CUST_ID</th>
<th>CUST_GENDER</th>
<th>CUST_MARITAL_STATUS</th>
<th>CUST_YEAR_OF_BIRTH</th>
<th>CUST_INCOME_LEVEL</th>
<th>CUST_CREDIT_LIMIT</th>
<th>EDUCATION</th>
<th>AFFINITY_CARD</th>
<th>HOUSEHOLD_SIZE</th>
<th>OCCUPATION</th>
<th>YRS_RESIDENCE</th>
<th>Y BOX_GAMES</th>
</tr>
</thead>
<tbody>
<tr>
<td>100.302</td>
<td>0</td>
<td>0.5908</td>
<td>0.4072</td>
<td>100.302</td>
<td>M</td>
<td>Married</td>
<td>1.975</td>
<td>J 150,000 - 240,999</td>
<td>11,000</td>
<td>10th</td>
<td>0</td>
<td>3</td>
<td>Crafts</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>100.302</td>
<td>1</td>
<td>0.4072</td>
<td>0.5928</td>
<td>100.302</td>
<td>M</td>
<td>Married</td>
<td>1.975</td>
<td>J 150,000 - 240,999</td>
<td>11,000</td>
<td>10th</td>
<td>0</td>
<td>3</td>
<td>Crafts</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>100.612</td>
<td>1</td>
<td>0.4625</td>
<td>0.4625</td>
<td>100.612</td>
<td>M</td>
<td>Married</td>
<td>1.974</td>
<td>H 150,000 - 169,999</td>
<td>7,000</td>
<td>Assoc-A</td>
<td>1</td>
<td>3</td>
<td>Sales</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>100.612</td>
<td>0</td>
<td>0.4625</td>
<td>0.5375</td>
<td>100.612</td>
<td>M</td>
<td>Married</td>
<td>1.974</td>
<td>H 150,000 - 169,999</td>
<td>7,000</td>
<td>Assoc-A</td>
<td>1</td>
<td>3</td>
<td>Sales</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>100.950</td>
<td>1</td>
<td>0.5928</td>
<td>0.4072</td>
<td>100.950</td>
<td>M</td>
<td>Married</td>
<td>1.958</td>
<td>G 130,000 - 149,999</td>
<td>5,000</td>
<td>HS-grad</td>
<td>1</td>
<td>3</td>
<td>Sales</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>100.950</td>
<td>0</td>
<td>0.5928</td>
<td>0.4072</td>
<td>100.950</td>
<td>M</td>
<td>Married</td>
<td>1.958</td>
<td>G 130,000 - 149,999</td>
<td>5,000</td>
<td>HS-grad</td>
<td>1</td>
<td>3</td>
<td>Sales</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>102.203</td>
<td>1</td>
<td>0.3735</td>
<td>0.4625</td>
<td>102.203</td>
<td>M</td>
<td>Married</td>
<td>1.966</td>
<td>I 170,000 - 199,999</td>
<td>9,000</td>
<td>Bach.</td>
<td>1</td>
<td>3</td>
<td>Sales</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>102.203</td>
<td>0</td>
<td>0.3735</td>
<td>0.4625</td>
<td>102.203</td>
<td>M</td>
<td>Married</td>
<td>1.966</td>
<td>I 170,000 - 199,999</td>
<td>9,000</td>
<td>Bach.</td>
<td>1</td>
<td>3</td>
<td>Sales</td>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>
How to Get Started with Oracle Machine Learning

Here is how you can get started with Oracle Machine Learning.

1. Request access to Oracle Machine Learning. Contact your Service Administrator to provide access to your Oracle Machine Learning account.

2. Access the Oracle Machine Learning account by using your credentials. In case you forget your password, then request the Administrator to reset it.

   **Note:**
   Once you receive your new password, you must change it immediately. Refer to the Oracle Machine Learning password policy for more information.

3. Once you log in for the first time, a workspace and project will be created for you. You can start creating your notebook and assign it to the default project and workspace. You can also create your own project and workspace.

Related Topics

- Password Policy

Accessing Oracle Machine Learning User Management Page

From Autonomous Data Warehouse Cloud you can access the Oracle Machine Learning **Manage Oracle ML Users** page.

To access Oracle Machine Learning **Manage Oracle ML Users** page:

1. Sign in to your Cloud Account and navigate to the **My Services** Dashboard.

2. Click the navigation menu icon ☰ in the top corner of the **My Services** Dashboard and then click **Autonomous Data Warehouse Cloud**.

3. Select a service and click the ☰ and select **Service Console**.

4. At the prompt, enter ADMIN for the username and enter the password for the ADMIN user.
Creating Projects and Workspaces

A project is a container for your notebooks, and a workspace is a container for your projects. You can own multiple projects in a workspace.

The initial workspace and the default project is created by the Oracle Machine Learning service automatically when you log in to Oracle Machine Learning for the first time. To create a new project and a workspace:

1. On the top right corner of Oracle Machine Learning home page, click the project workspace drop-down list. The project name and the workspace, in which the project resides, are displayed here. In this screenshot, the project name is Project A, and the workspace name is Admin. If a default project exists, then the default project name is displayed here. To choose a different project, click Select Project.

2. To create a new project, click New Project.
   The Create Project dialog box opens.
3. In the Name field, provide a name for your project.
4. In the Comments field, enter comments, if any.
5. In the Select Workspace field, select a workspace from the drop-down list. Your project is assigned to the selected workspace. If you want to create a new workspace, then click New Workspace.
6. In the Create Workspace dialog box, enter a name for the workspace in the Name field.
7. In the Comments field, enter comments, if any.
8. Click OK.
   This creates your workspace and navigates back to the Create Project dialog box. The project that you are creating is now assigned to the newly created workspace.
9. Click OK.
Oracle’s Machine Learning/Advanced Analytics Platforms

Machine Learning Algorithms Embedded in the Data Management Platforms

“Analytics Producers”

Data Scientists, R Users, Citizen Data Scientists

New Zeppelin notebook based UI for data scientists collaborating and sharing ML analytical methodologies in Clouds

Big Data Cloud Service

“Oracle Machine Learning” Big Data Cloud

ORA AH — Machine Learning Algorithms
Statistical Functions + R Integration
for Scalable, Parallel, Distributed Execution

Database Cloud

“Oracle Machine Learning” Database Edition

Machine Learning Algorithms,
Statistical Functions + R Integration
for Scalable, Parallel, Distributed, in-DB Execution

Oracle’s Machine Learning/Advanced Analytics Platforms

Machine Learning Algorithms Embedded in the Data Management Platforms

“Analytics Producers”

Data Scientists, R Users, Citizen Data Scientists

New Zeppelin notebook based UI for data scientists collaborating and sharing ML analytical methodologies in Clouds

Big Data Cloud Service

“Oracle Machine Learning” Big Data Cloud

ORA AH — Machine Learning Algorithms
Statistical Functions + R Integration
for Scalable, Parallel, Distributed Execution

Database Cloud

“Oracle Machine Learning” Database Edition

Machine Learning Algorithms,
Statistical Functions + R Integration
for Scalable, Parallel, Distributed, in-DB Execution

Oracle’s Machine Learning/Advanced Analytics Platforms

Machine Learning Algorithms Embedded in the Data Management Platforms

“Analytics Producers”

Data Scientists, R Users, Citizen Data Scientists

New Zeppelin notebook based UI for data scientists collaborating and sharing ML analytical methodologies in Clouds

Big Data Cloud Service

“Oracle Machine Learning” Big Data Cloud

ORA AH — Machine Learning Algorithms
Statistical Functions + R Integration
for Scalable, Parallel, Distributed Execution

Database Cloud

“Oracle Machine Learning” Database Edition

Machine Learning Algorithms,
Statistical Functions + R Integration
for Scalable, Parallel, Distributed, in-DB Execution

Oracle’s Machine Learning/Advanced Analytics Platforms

Machine Learning Algorithms Embedded in the Data Management Platforms

“Analytics Producers”

Data Scientists, R Users, Citizen Data Scientists

New Zeppelin notebook based UI for data scientists collaborating and sharing ML analytical methodologies in Clouds

Big Data Cloud Service

“Oracle Machine Learning” Big Data Cloud

ORA AH — Machine Learning Algorithms
Statistical Functions + R Integration
for Scalable, Parallel, Distributed Execution

Database Cloud

“Oracle Machine Learning” Database Edition

Machine Learning Algorithms,
Statistical Functions + R Integration
for Scalable, Parallel, Distributed, in-DB Execution

Copyright © 2018, Oracle and/or its affiliates. All rights reserved.
Oracle Machine Learning and Advanced Analytics

Strategy and Road Map

• Support multiple data platforms, analytical engines, languages, UIs and deployment strategies

GUI

Data Miner, RStudio

Notebooks

SQL

ML Algorithms

Common core, parallel, distributed

R, Python, etc.

Big Data / Big Data Cloud

Relational

Oracle Database Cloud
More Information on Oracle’s ML/AA Functionality
Oracle’s Machine Learning/Advanced Analytics

Fastest Way to Deliver Scalable Enterprise-wide Predictive Analytics

Key Features

- Parallel, scalable machine learning algorithms and R integration
- In-Database + Hadoop—Don’t move the data
- Data analysts, data scientists & developers
- Drag and drop workflow, R and SQL APIs
- Extends data management into powerful advanced/predictive analytics platform
- Enables enterprise predictive analytics deployment + applications
Multiple Data Scientist User Roles Supported
Oracle’s Machine Learning/Advanced Analytics

DBAs

Application Developers

R Users, Data Scientists

Data Analysts, Citizen Data Scientists

Additional relevant data and “engineered” features:
- Sensor data, Text, unstructured data, transactional data, spatial data, etc.
- Historical data

Historical or current data to be "scored" for predictions

Oracle Database 12c
Multiple Data Scientist User Roles Supported

Oracle’s Machine Learning/Advanced Analytics

- R Users, Data Scientists
- OML SQL Notebook Users
- Data Analyst, Citizen Data Scientists

Additional relevant data and “engineered features”

Sensor data, Text, unstructured data, transactional data, spatial data, etc.

Historical or Current Data to be “scored”

Predictions & Insights

Oracle Database 12c

Copyright © 2018, Oracle and/or its affiliates. All rights reserved.
Manage and Analyze All Your Data

Data Scientists, R Users, Citizen Data Scientists

Architecturally, Many Options and Flexibility

SQL / R

Boil down the Data Lake

Big Data SQL / R

Object Store  NoSQL  ORACLE  kafka

“Engineered Features”

– Derived attributes that reflect domain knowledge—key to best models e.g:
  • Counts
  • Totals
  • Changes over time

**Classification**
- Naïve Bayes
- Logistic Regression (GLM)
- Decision Tree
- Random Forest
- Neural Network
- Support Vector Machine
- Explicit Semantic Analysis

**CLUSTERING**
- Hierarchical K-Means
- Hierarchical O-Cluster
- Expectation Maximization (EM)

**ANOMALY DETECTION**
- One-Class SVM

**TIME SERIES**
- Holt-Winters, Regular & Irregular, with and w/o trends & seasonal
- Single, Double Exp Smoothing

**REGRESSION**
- Linear Model
- Generalized Linear Model
- Support Vector Machine (SVM)
- Stepwise Linear regression
- Neural Network
- LASSO

**ATTRIBUTE IMPORTANCE**
- Minimum Description Length
- Principal Comp Analysis (PCA)
- Unsupervised Pair-wise KL Div
- CUR decomposition for row & AI

**ASSOCIATION RULES**
- A priori/ market basket

**PREDICTIVE QUERIES**
- Predict, cluster, detect, features

**SQL ANALYTICS**
- SQL Windows, SQL Patterns, SQL Aggregates

**FEATURE EXTRACTION**
- Principal Comp Analysis (PCA)
- Non-negative Matrix Factorization
- Singular Value Decomposition (SVD)
- Explicit Semantic Analysis (ESA)

**TEXT MINING SUPPORT**
- Algorithms support text type
- Tokenization and theme extraction
- Explicit Semantic Analysis (ESA) for document similarity

**STATISTICAL FUNCTIONS**
- Basic statistics: min, max, median, stdev, t-test, F-test, Pearson’s, Chi-Sq, ANOVA, etc.

**R PACKAGES**
- CRAN R Algorithm Packages through Embedded R Execution
- Spark MLlib algorithm integration

**EXPORTABLE ML MODELS**
- C and Java code for deployment

---

* OAA (Oracle Data Mining + Oracle R Enterprise) and ORAAH combined
* OAA includes support for Partitioned Models, Transactional, Unstructured, Geo-spatial, Graph data, etc.

Copyright © 2018, Oracle and/or its affiliates. All rights reserved.
Oracle’s Machine Learning/Advanced Analytics
Fastest Way to Deliver Enterprise-wide Predictive Analytics

Major Benefits

- Data remains in Database & Hadoop
  - Model building and scoring occur in-database
  - Use R packages with data-parallel invocations
- Leverage investment in Oracle IT
  - Eliminate data duplication
  - Eliminate separate analytical servers
- Deliver enterprise-wide applications
  - GUI for ML/Predictive Analytics & code gen
  - R interface leverages database as HPC engine

Traditional Analytics
- Data Import
- Data Mining
- Model “Scoring”
- Data Prep. & Transformation
- Data Extraction
- Hours, Days or Weeks

Oracle Advanced Analytics
- Data Import
- Data Mining
- Model Building
- Data Prep & Transformation
- Data Extraction
- Secs, Mins or Hours

Savings
Oracle Advanced Analytics 12.2
Model Build Time Performance

<table>
<thead>
<tr>
<th>OAA 12.2 Algorithms</th>
<th>Rows (Ms)</th>
<th>T7-4 (Sparc &amp; Solaris)</th>
<th>X5-4 (Intel and Linux)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attributes Importance</td>
<td>640</td>
<td>28s / 512</td>
<td>44s / 72</td>
</tr>
<tr>
<td>K Means Clustering</td>
<td>640</td>
<td>161s / 256</td>
<td>268s / 144</td>
</tr>
<tr>
<td>Expectation Maximization</td>
<td>159</td>
<td>455s / 512</td>
<td>588s / 144</td>
</tr>
<tr>
<td>Naive Bayes Classification</td>
<td>320</td>
<td>17s / 256</td>
<td>23s / 72</td>
</tr>
<tr>
<td>GLM Classification</td>
<td>640</td>
<td>154s / 512</td>
<td>363s / 144</td>
</tr>
</tbody>
</table>

GLM Regression

In 24 hours, could build new predictive models for entire Support Vector United States Population, for 400 attributes, 4 times!

The way to read their results is that they compare 2 chips: X5 (Intel and Linux) and T7 (Sparc and Solaris). They are measuring scalability (time in seconds) with increase degree of parallelism (dop). The data also has high cardinality categorical columns which translates in high mining attributes (when algorithms require explosion). There are no comparisons to 12.1 and it is fair to say that the 12.1 algorithms could not run on data of this size.
Fraud Prediction Demo

Automated In-DB Analytical Methodology

drop table CLAIMS_SET;
exec dbms_data_mining.drop_model('CLAIMSMODEL');
create table CLAIMS_SET (setting_name varchar2(30), setting_value varchar2(4000));
insert into CLAIMS_SET values ('ALGO_NAME','ALGO_SUPPORT_VECTOR_MACHINES');
insert into CLAIMS_SET values ('PREP_AUTO','ON');
commit;

begin
  dbms_data_mining.create_model('CLAIMSMODEL', 'CLASSIFICATION', 'CLAIMS', 'POLICYNUMBER', null, 'CLAIMS_SET');
end;
/

-- Top 5 most suspicious fraud policy holder claims
select * from
(select POLICYNUMBER, round(prob_fraud*100,2) percent_fraud,
  rank() over (order by prob_fraud desc) rnk from
(select POLICYNUMBER, prediction_probability(CLAIMSMODEL, '0' using *) prob_fraud
from CLAIMS
where PASTNUMBEROFCLAIMS in ('2to4', 'morethan4'))
where rnk <= 5
order by percent_fraud desc;

Automated Monthly “Application”!
Just add:
Create View CLAIMS2_30
As
Select * from CLAIMS2
Where mydate > SYSDATE – 30

Time measure: set timing on;
Oracle Advanced Analytics
Real-Time Scoring, Predictions and Recommendations

• On-the-fly, single record apply with new data (e.g. from call center)

```sql
Select prediction_probability(CLAS_DT_1_15, 'Yes'
    USING 7800 as bank_funds, 125 as checking_amount, 20 as credit_balance, 55 as age, 'Married' as marital_status,
    250 as MONEY_MONTLY_OVERDRAWN, 1 as house_ownership)
from dual;
```

Likelihood to respond:
# Build Predictive Models on an Attribute

Oracle's Machine Learning Accelerates New Possibilities

## Machine Learning Model

$$\text{Function}(X_1, X_2, \ldots, X) \rightarrow Y \ (\text{LTV\_BIN})\; \text{Probability}$$

<table>
<thead>
<tr>
<th>CUST_ID</th>
<th>AGE</th>
<th>SEX</th>
<th>MARITAL_STATUS</th>
<th>N_TRANS_ATM</th>
<th>LTV</th>
<th>CHECKING_AMOUNT</th>
<th>BANK_FUNDS</th>
<th>SALARY</th>
<th>HOUSE_OWNERSHIP</th>
<th>PROFESSION</th>
<th>LTV_BIN</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU8817</td>
<td>46.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>34,040.000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,760</td>
<td>2.0000</td>
<td>Cashier</td>
<td>VERY HIGH</td>
<td>91</td>
</tr>
<tr>
<td>CU7115</td>
<td>27.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>20,843,250</td>
<td>25.0000</td>
<td>0.0000</td>
<td>58,573</td>
<td>1.0000</td>
<td>PROF-14</td>
<td>MEDIUM</td>
<td>77</td>
</tr>
<tr>
<td>CU7117</td>
<td>30.0000</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>28,306,500</td>
<td>349.0000</td>
<td>6,300.000</td>
<td>65,226</td>
<td>2.0000</td>
<td>Administrator</td>
<td>HIGH</td>
<td>64</td>
</tr>
<tr>
<td>CU7118</td>
<td>33.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>1.0000</td>
<td>26,680,250</td>
<td>25.0000</td>
<td>0.0000</td>
<td>68,721</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7120</td>
<td>21.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>3.0000</td>
<td>22,012,000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>55,948</td>
<td>1.0000</td>
<td>Professor</td>
<td>HIGH</td>
<td>52</td>
</tr>
<tr>
<td>CU7121</td>
<td>32.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>20,604,500</td>
<td>999.0000</td>
<td>20,100.000</td>
<td>66,818</td>
<td>1.0000</td>
<td>PROF-36</td>
<td>MEDIUM</td>
<td>93</td>
</tr>
<tr>
<td>CU7123</td>
<td>35.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>22,330,500</td>
<td>92.0000</td>
<td>8,200.000</td>
<td>61,322</td>
<td>1.0000</td>
<td>Clerk</td>
<td>HIGH</td>
<td>83</td>
</tr>
<tr>
<td>CU7124</td>
<td>44.0000</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>23,085,500</td>
<td>25.0000</td>
<td>2,000.000</td>
<td>66,742</td>
<td>1.0000</td>
<td>PROF-1</td>
<td>HIGH</td>
<td>54</td>
</tr>
<tr>
<td>CU7125</td>
<td>63.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>1.0000</td>
<td>25,383,000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,332</td>
<td>1.0000</td>
<td>Water/Waitress</td>
<td>HIGH</td>
<td>74</td>
</tr>
<tr>
<td>CU7127</td>
<td>58.0000</td>
<td>M</td>
<td>WIDOWED</td>
<td>2.0000</td>
<td>29,106,000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,724</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>68</td>
</tr>
<tr>
<td>CU7128</td>
<td>26.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>15,084,750</td>
<td>708.0000</td>
<td>1,100,000</td>
<td>65,939</td>
<td>0.0000</td>
<td>Programmer</td>
<td>MEDIUM</td>
<td>55</td>
</tr>
<tr>
<td>CU7129</td>
<td>28.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>0.0000</td>
<td>30,251,750</td>
<td>25.0000</td>
<td>0.0000</td>
<td>66,207</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7131</td>
<td>47.0000</td>
<td>F</td>
<td>DIVORCED</td>
<td>4.0000</td>
<td>25,222,750</td>
<td>1,186.000</td>
<td>2,600.000</td>
<td>66,091</td>
<td>1.0000</td>
<td>School Teacher</td>
<td>HIGH</td>
<td>93</td>
</tr>
<tr>
<td>CU7132</td>
<td>39.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>32,064,750</td>
<td>25.0000</td>
<td>0.0000</td>
<td>66,659</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>98</td>
</tr>
<tr>
<td>CU7133</td>
<td>26.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>3.0000</td>
<td>15,426,250</td>
<td>25.0000</td>
<td>0.0000</td>
<td>61,301</td>
<td>1.0000</td>
<td>Fireman</td>
<td>MEDIUM</td>
<td>92</td>
</tr>
<tr>
<td>CU7134</td>
<td>38.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>20,695,250</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,381</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>90</td>
</tr>
<tr>
<td>CU7135</td>
<td>42.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>21,000,750</td>
<td>25.0000</td>
<td>0.0000</td>
<td>66,303</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>76</td>
</tr>
<tr>
<td>CU7136</td>
<td>36.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>22,064,750</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,859</td>
<td>0.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>92</td>
</tr>
<tr>
<td>CU7137</td>
<td>23.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>20,237,250</td>
<td>25.0000</td>
<td>0.0000</td>
<td>67,749</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>65</td>
</tr>
<tr>
<td>CU7138</td>
<td>38.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>5.0000</td>
<td>24,152,500</td>
<td>25.0000</td>
<td>2,150,000</td>
<td>67,410</td>
<td>1.0000</td>
<td>Veterinarian</td>
<td>HIGH</td>
<td>93</td>
</tr>
</tbody>
</table>
## Build Predictive Models on an Attribute

Oracle’s Machine Learning Accelerates New Possibilities

### Machine Learning Model

\[ \text{Function}(X_1, X_2, ..., X) \]

\[ \Rightarrow Y (\text{LTV\_BIN}); \quad \text{Probability} \]

<table>
<thead>
<tr>
<th>CUST_ID</th>
<th>AGE</th>
<th>SEX</th>
<th>MARITAL_STATUS</th>
<th>N_TRANS_ATM</th>
<th>LTV</th>
<th>CHECKING_AMOUNT</th>
<th>BANK_FUNDS</th>
<th>SALARY</th>
<th>HOUSE_OWNERSHIP</th>
<th>PROFESSION</th>
<th>LTV_BIN</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU8817</td>
<td>46.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>34,040.000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,760...</td>
<td>2.0000</td>
<td>Cashier</td>
<td>VERY HIGH</td>
<td>91</td>
</tr>
<tr>
<td>CU7115</td>
<td>27.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>20,843.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>56,739...</td>
<td>1.0000</td>
<td>PROF-14</td>
<td>MEDIUM</td>
<td>77</td>
</tr>
<tr>
<td>CU7117</td>
<td>30.0000</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>28,306.5000</td>
<td>349.0000</td>
<td>6,300.000</td>
<td>65,225...</td>
<td>2.0000</td>
<td>Adminstrator</td>
<td>HIGH</td>
<td>64</td>
</tr>
<tr>
<td>CU7118</td>
<td>32.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>1.0000</td>
<td>26,480.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>68,721...</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7120</td>
<td>21.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>3.0000</td>
<td>22,012.0000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>55,648...</td>
<td>1.0000</td>
<td>Professor</td>
<td>HIGH</td>
<td>52</td>
</tr>
<tr>
<td>CU7121</td>
<td>32.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>29,904.5000</td>
<td>999.0000</td>
<td>20,100.000</td>
<td>65,818...</td>
<td>1.0000</td>
<td>PROF-36</td>
<td>MEDIUM</td>
<td>93</td>
</tr>
<tr>
<td>CU7123</td>
<td>35.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>22,330.5000</td>
<td>92.0000</td>
<td>8,200.000</td>
<td>61,322...</td>
<td>1.0000</td>
<td>Cook</td>
<td>HIGH</td>
<td>83</td>
</tr>
<tr>
<td>CU7124</td>
<td>44.0000</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>23,085.5000</td>
<td>25.0000</td>
<td>2,000.000</td>
<td>64,742...</td>
<td>1.0000</td>
<td>PROF-1</td>
<td>HIGH</td>
<td>54</td>
</tr>
<tr>
<td>CU7125</td>
<td>63.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>1.0000</td>
<td>25,383.0000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>62,332...</td>
<td>1.0000</td>
<td>Water/Waitress</td>
<td>HIGH</td>
<td>74</td>
</tr>
<tr>
<td>CU7127</td>
<td>58.0000</td>
<td>M</td>
<td>WIDOWED</td>
<td>2.0000</td>
<td>29,106.0000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,224...</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>68</td>
</tr>
<tr>
<td>CU7128</td>
<td>26.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>15,084.7500</td>
<td>708.0000</td>
<td>1,100.000</td>
<td>65,939...</td>
<td>0.0000</td>
<td>Programmer</td>
<td>MEDIUM</td>
<td>55</td>
</tr>
<tr>
<td>CU7129</td>
<td>28.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>0.0000</td>
<td>30,251.7500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>66,207...</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7131</td>
<td>47.0000</td>
<td>F</td>
<td>DIVORCED</td>
<td>4.0000</td>
<td>25,222.7500</td>
<td>1,186.0000</td>
<td>2,600.000</td>
<td>66,091...</td>
<td>1.0000</td>
<td>School Teacher</td>
<td>HIGH</td>
<td>93</td>
</tr>
<tr>
<td>CU7132</td>
<td>39.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>32,064.7500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>66,659...</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>98</td>
</tr>
<tr>
<td>CU7133</td>
<td>26.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>3.0000</td>
<td>15,425.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>61,301...</td>
<td>1.0000</td>
<td>Fireman</td>
<td>MEDIUM</td>
<td>89</td>
</tr>
<tr>
<td>CU7134</td>
<td>38.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>20,695.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>60,381...</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>90</td>
</tr>
<tr>
<td>CU7135</td>
<td>42.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>21,000.7500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>66,303...</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>76</td>
</tr>
<tr>
<td>CU7138</td>
<td>36.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>22,064.7500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,859...</td>
<td>0.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>92</td>
</tr>
<tr>
<td>CU7137</td>
<td>23.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>20,237.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>65,749...</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>65</td>
</tr>
<tr>
<td>CU7138</td>
<td>38.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>5.0000</td>
<td>24,152.5000</td>
<td>25.0000</td>
<td>2,150.000</td>
<td>67,410...</td>
<td>1.0000</td>
<td>Veterinarian</td>
<td>HIGH</td>
<td>93</td>
</tr>
</tbody>
</table>
Build Predictive Models on an Attribute
Oracle’s Machine Learning Accelerates New Possibilities

Machine Learning Models \( \rightarrow \) Function\((X_1, X_2, \ldots, X)\) \( \rightarrow \) Y2 (BankFunds)

\[
Y (LTV\_BIN);\quad \text{Probability}
\]

<table>
<thead>
<tr>
<th>CUST_ID</th>
<th>AGE</th>
<th>SEX</th>
<th>MARITAL_STATUS</th>
<th>N_TRANS_ATM</th>
<th>LTV</th>
<th>CHECKING_AMOUNT</th>
<th>BANK_FUNDS</th>
<th>P(BankFunds)</th>
<th>SALARY</th>
<th>HOUSE_OWNERSHIP</th>
<th>PROFESSION</th>
<th>LTV_BIN</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU8817</td>
<td>46.000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>34,040.000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>4,500,000</td>
<td>65,760</td>
<td>2.0000</td>
<td>Cashier</td>
<td>VERY HIGH</td>
<td>91</td>
</tr>
<tr>
<td>CU7115</td>
<td>35.000</td>
<td>M</td>
<td>MARRIED</td>
<td>5.0000</td>
<td>20,843.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>1,500,000</td>
<td>58,573</td>
<td>1.0000</td>
<td>PROF-14</td>
<td>MEDIUM</td>
<td>77</td>
</tr>
<tr>
<td>CU7117</td>
<td>30.000</td>
<td>F</td>
<td>MARRIED</td>
<td>5.0000</td>
<td>28,306.5000</td>
<td>549.0000</td>
<td>6,300.000</td>
<td>35,000,000</td>
<td>65,225</td>
<td>2.0000</td>
<td>Administrator</td>
<td>HIGH</td>
<td>64</td>
</tr>
<tr>
<td>CU7118</td>
<td>53.000</td>
<td>M</td>
<td>MARRIED</td>
<td>1.0000</td>
<td>26,480.000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>150,000,000</td>
<td>68,721</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7120</td>
<td>32.000</td>
<td>M</td>
<td>MARRIED</td>
<td>3.0000</td>
<td>22,012.0000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>-5,000,000</td>
<td>57,948</td>
<td>1.0000</td>
<td>Professor</td>
<td>HIGH</td>
<td>83</td>
</tr>
<tr>
<td>CU7121</td>
<td>32.000</td>
<td>F</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>20,904.500</td>
<td>999.0000</td>
<td>20,10000</td>
<td>210,000</td>
<td>65,818</td>
<td>1.0000</td>
<td>PROF-36</td>
<td>MEDIUM</td>
<td>93</td>
</tr>
<tr>
<td>CU7123</td>
<td>35.000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>22,330.5000</td>
<td>92.0000</td>
<td>8,200.000</td>
<td>8,500,000</td>
<td>63,322</td>
<td>1.0000</td>
<td>Coach</td>
<td>HIGH</td>
<td>83</td>
</tr>
<tr>
<td>CU7124</td>
<td>44.000</td>
<td>F</td>
<td>MARRIED</td>
<td>5.0000</td>
<td>23,085.5000</td>
<td>25.0000</td>
<td>2,000.000</td>
<td>2,500,000</td>
<td>66,742</td>
<td>1.0000</td>
<td>PROF-1</td>
<td>HIGH</td>
<td>54</td>
</tr>
<tr>
<td>CU7125</td>
<td>63.000</td>
<td>M</td>
<td>MARRIED</td>
<td>1.0000</td>
<td>25,383.000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>25,000</td>
<td>65,332</td>
<td>1.0000</td>
<td>Waiter/Waitress</td>
<td>HIGH</td>
<td>74</td>
</tr>
<tr>
<td>CU7127</td>
<td>59.000</td>
<td>M</td>
<td>WIDOWED</td>
<td>2.0000</td>
<td>29,106.000</td>
<td>25.0000</td>
<td>0.0000</td>
<td>11,000</td>
<td>68,224</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>68</td>
</tr>
<tr>
<td>CU7128</td>
<td>26.000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>15,084.7500</td>
<td>708.0000</td>
<td>1,100.000</td>
<td>1,200,000</td>
<td>65,939</td>
<td>0.0000</td>
<td>Programmer</td>
<td>MEDIUM</td>
<td>55</td>
</tr>
<tr>
<td>CU7129</td>
<td>28.000</td>
<td>M</td>
<td>MARRIED</td>
<td>0.0000</td>
<td>30,251.7500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>17,000</td>
<td>66,207</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7131</td>
<td>47.000</td>
<td>F</td>
<td>DIVORCED</td>
<td>4.0000</td>
<td>25,222.7500</td>
<td>1,186.000</td>
<td>2,600.000</td>
<td>2,750,000</td>
<td>66,911</td>
<td>1.0000</td>
<td>School Teacher</td>
<td>HIGH</td>
<td>93</td>
</tr>
<tr>
<td>CU7132</td>
<td>39.000</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>32,064.7500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>0</td>
<td>66,659</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>98</td>
</tr>
<tr>
<td>CU7133</td>
<td>26.000</td>
<td>F</td>
<td>MARRIED</td>
<td>3.0000</td>
<td>15,425.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>-4,500,000</td>
<td>61,301</td>
<td>1.0000</td>
<td>Fireman</td>
<td>MEDIUM</td>
<td>89</td>
</tr>
<tr>
<td>CU7134</td>
<td>38.000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>20,695.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>9,000</td>
<td>63,581</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>90</td>
</tr>
<tr>
<td>CU7135</td>
<td>42.000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>21,900.7500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>6,500</td>
<td>66,803</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>76</td>
</tr>
<tr>
<td>CU7136</td>
<td>36.000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>22,644.7500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>12,000</td>
<td>66,959</td>
<td>0.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>92</td>
</tr>
<tr>
<td>CU7137</td>
<td>33.000</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>20,237.2500</td>
<td>25.0000</td>
<td>0.0000</td>
<td>7,500</td>
<td>67,749</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>65</td>
</tr>
<tr>
<td>CU7138</td>
<td>38.000</td>
<td>F</td>
<td>MARRIED</td>
<td>5.0000</td>
<td>24,152.5000</td>
<td>25.0000</td>
<td>2,150.000</td>
<td>2,200,000</td>
<td>67,910</td>
<td>1.0000</td>
<td>Veterinarian</td>
<td>HIGH</td>
<td>93</td>
</tr>
</tbody>
</table>
Build Predictive Models on an Attribute
Oracle’s Machine Learning Accelerates New Possibilities

Machine Learning Models ➔ Function(\(X_1, X_2, \ldots, X\)) ➔ \(Y_2\) (BankFunds)

\[ Y_2 (\text{BankFunds}) = \text{Function}(X_1, X_2, \ldots, X) \]

\[ Y (\text{LTV BIN}) \] Probability

<table>
<thead>
<tr>
<th>CUST_ID</th>
<th>AGE</th>
<th>SEX</th>
<th>MARITAL_STATUS</th>
<th>N_TRANS_ATM</th>
<th>LTV</th>
<th>CHECKING_AMOUNT</th>
<th>BANK_FUNDS (PV)</th>
<th>SALARY</th>
<th>HOUSE_OWNERSHIP</th>
<th>PROFESSION</th>
<th>LTिव_BIN</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU8817</td>
<td>46.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>34,040.000</td>
<td>25.000</td>
<td>0.000</td>
<td>65,760</td>
<td>2.0000</td>
<td>Cashier</td>
<td>VERY HIGH</td>
<td>91</td>
</tr>
<tr>
<td>CU7115</td>
<td>27.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>20,843.2500</td>
<td>25.000</td>
<td>0.0000</td>
<td>58,573</td>
<td>1.0000</td>
<td>PROF-14</td>
<td>MEDIUM</td>
<td>77</td>
</tr>
<tr>
<td>CU7117</td>
<td>30.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>28,306.5000</td>
<td>349.000</td>
<td>0.0000</td>
<td>65,226</td>
<td>2.0000</td>
<td>Administrator</td>
<td>HIGH</td>
<td>64</td>
</tr>
<tr>
<td>CU7118</td>
<td>33.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>1.0000</td>
<td>26,480.2500</td>
<td>25.000</td>
<td>0.0000</td>
<td>68,721</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7120</td>
<td>21.0000</td>
<td>M</td>
<td>DIVORCED</td>
<td>3.0000</td>
<td>22,012.0000</td>
<td>25.000</td>
<td>0.0000</td>
<td>57,948</td>
<td>1.0000</td>
<td>Professor</td>
<td>HIGH</td>
<td>84</td>
</tr>
<tr>
<td>CU7121</td>
<td>32.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>20,904.5000</td>
<td>999.000</td>
<td>0.0000</td>
<td>66,818</td>
<td>1.0000</td>
<td>PROF-38</td>
<td>MEDIUM</td>
<td>93</td>
</tr>
<tr>
<td>CU7123</td>
<td>35.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>22,330.5000</td>
<td>92.000</td>
<td>8,200.000</td>
<td>62,832</td>
<td>1.0000</td>
<td>Closed</td>
<td>HIGH</td>
<td>83</td>
</tr>
<tr>
<td>CU7124</td>
<td>44.0000</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>23,085.5000</td>
<td>25.000</td>
<td>2,000.000</td>
<td>66,742</td>
<td>1.0000</td>
<td>PROF-1</td>
<td>HIGH</td>
<td>54</td>
</tr>
<tr>
<td>CU7125</td>
<td>63.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>1.0000</td>
<td>25,383.0000</td>
<td>25.000</td>
<td>0.0000</td>
<td>62,332</td>
<td>1.0000</td>
<td>Water/Waitress</td>
<td>HIGH</td>
<td>74</td>
</tr>
<tr>
<td>CU7127</td>
<td>58.0000</td>
<td>M</td>
<td>WIDOWED</td>
<td>2.0000</td>
<td>29,106.0000</td>
<td>25.000</td>
<td>0.0000</td>
<td>68,224</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>68</td>
</tr>
<tr>
<td>CU7128</td>
<td>26.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>15,084.7500</td>
<td>7082.000</td>
<td>1,100.000</td>
<td>65,939</td>
<td>0.0000</td>
<td>Programmer/</td>
<td>MEDIUM</td>
<td>55</td>
</tr>
<tr>
<td>CU7129</td>
<td>28.0000</td>
<td>M</td>
<td>MARRIED</td>
<td>0.0000</td>
<td>30,251.7500</td>
<td>25.000</td>
<td>0.0000</td>
<td>66,297</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7132</td>
<td>47.0000</td>
<td>F</td>
<td>DIVORCED</td>
<td>4.0000</td>
<td>25,222.7500</td>
<td>1,186.2500</td>
<td>2,600.000</td>
<td>66,591</td>
<td>1.0000</td>
<td>School Teacher</td>
<td>HIGH</td>
<td>93</td>
</tr>
<tr>
<td>CU7132</td>
<td>39.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>32,064.7500</td>
<td>25.000</td>
<td>0.0000</td>
<td>66,629</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>98</td>
</tr>
<tr>
<td>CU7133</td>
<td>26.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>3.0000</td>
<td>15,425.2500</td>
<td>25.000</td>
<td>0.0000</td>
<td>61,301</td>
<td>1.0000</td>
<td>Fireman</td>
<td>MEDIUM</td>
<td>89</td>
</tr>
<tr>
<td>CU7134</td>
<td>38.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>20,695.2500</td>
<td>25.000</td>
<td>0.0000</td>
<td>63,581</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>90</td>
</tr>
<tr>
<td>CU7135</td>
<td>42.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>21,900.7500</td>
<td>25.000</td>
<td>0.0000</td>
<td>66,803</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>76</td>
</tr>
<tr>
<td>CU7136</td>
<td>36.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>22,064.7500</td>
<td>25.000</td>
<td>0.0000</td>
<td>69,859</td>
<td>0.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>92</td>
</tr>
<tr>
<td>CU7137</td>
<td>23.0000</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>20,237.2500</td>
<td>25.000</td>
<td>0.0000</td>
<td>67,749</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>65</td>
</tr>
<tr>
<td>CU7138</td>
<td>38.0000</td>
<td>F</td>
<td>MARRIED</td>
<td>5.0000</td>
<td>24,152.5000</td>
<td>25.000</td>
<td>2,150.000</td>
<td>67,910</td>
<td>1.0000</td>
<td>Veterinarian</td>
<td>HIGH</td>
<td>93</td>
</tr>
</tbody>
</table>
Oracle Data Miner “Workflow” UI
Easy to use for “Citizen Data Scientist”; Fast to Deploy via SQL and PL/SQL Scripts

- SQL Developer Extension
- Easy to use to define analytical methodologies that can be shared
- Workflow API and generates SQL code for deployment

Step 1. “Citizen Data Scientist” builds, tests, applies ML methodologies using DS IDEs (ODMr, RStudio)

Step 2. Insights & Predictions in Database for OAC/DV user additional Viz/Analytics
### Importance Value, Rank by Attribute Name

<table>
<thead>
<tr>
<th>Attribute Name</th>
<th>Importance Value</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consumer_Findex_Score</td>
<td>0.74</td>
<td>1</td>
</tr>
<tr>
<td>Customer_DMG_Segment</td>
<td>0.54</td>
<td>3</td>
</tr>
<tr>
<td>Customer_Lifetime_Vaule</td>
<td>0.08</td>
<td>13</td>
</tr>
<tr>
<td>Customer_Value_Segment</td>
<td>0.40</td>
<td>6</td>
</tr>
<tr>
<td>Delinquency_Status</td>
<td>0.23</td>
<td>11</td>
</tr>
<tr>
<td>Education_Level</td>
<td>0.17</td>
<td>12</td>
</tr>
<tr>
<td>Family_Size</td>
<td>0.04</td>
<td>14</td>
</tr>
<tr>
<td>Highest_Credit_Card_Limit</td>
<td>0.42</td>
<td>4</td>
</tr>
<tr>
<td>Income</td>
<td>0.42</td>
<td>5</td>
</tr>
<tr>
<td>Max_CC_Spent_Amount</td>
<td>0.25</td>
<td>9</td>
</tr>
<tr>
<td>Max_CC_Spent_Amount_Prev</td>
<td>0.27</td>
<td>8</td>
</tr>
<tr>
<td>New_Bankruptcy</td>
<td>0.00</td>
<td>16</td>
</tr>
<tr>
<td>Number_of_Collections</td>
<td>0.00</td>
<td>17</td>
</tr>
<tr>
<td>Number_of_Liabilities</td>
<td>0.04</td>
<td>15</td>
</tr>
<tr>
<td>Occupation</td>
<td>0.24</td>
<td>10</td>
</tr>
</tbody>
</table>

### Attribute Name by Importance Value

- Consumer_Findex_Score
- Wealth
- Highest_Credit_Card_Limit
- Income
- Customer_Value_Segment
- Residential_Status
- Max_CC_Spent_Amount
- OCCUPATION
- Delinquency_Status
- Education_Level
- Customer_Lifetime_Value
- Family_Size
- New_Bankruptcy
- Number_of_Liabilities
- Number_of_Collections
The Core Ingredients of Good Machine Learning

Domain Knowledge + Data

Machine Learning Algorithms

Insights, Predictions

Copyright © 2018, Oracle and/or its affiliates. All rights reserved.
Most Important Factor in Machine Learning? **Deployment!!**

A “Thinking” Database
Summary & Next Steps
Oracle’s Machine Learning & Advanced Analytics Data Management Platforms

Summary

• Machine learning, predictive analytics & “AI” have become must-have requirements

• Enterprises whose data science teams most rapidly extract predictions and insights win

• Separate islands for data management and for data science don’t work

• Evolve towards combined data management + advanced analytics environ that can analyze data, perform machine learning and essentially to “think”

• “Operationalize” ML methodologies and discovered insights & predictions thru organizations for process automation and customer behavior anticipation
Getting Started—Oracle ML/AA Resources & Links

**Oracle Advanced Analytics Overview Information**
- Oracle's Machine Learning and Advanced Analytics 12.2c and Oracle Data Miner 4.2 New Features presentation
- Oracle Advanced Analytics Public Customer References
- Oracle’s Machine Learning and Advanced Analytics Data Management Platforms white paper on OTN
- Oracle INTERNAL ONLY OAA Product Management Wiki and Beehive Workspace (contains latest presentations, demos, product, etc. information)

**YouTube recorded Oracle Advanced Analytics Presentations and Demos, White Papers**
- Oracle's Machine Learning & Advanced Analytics 12.2 & Oracle Data Miner 4.2 New Features YouTube video
- Library of YouTube Movies on Oracle Advanced Analytics, Data Mining, Machine Learning (7+ “live” Demos e.g. Oracle Data Miner 4.0 New Features, Retail, Fraud, Loyalty, Overview, etc.)
- Overview YouTube video of Oracle’s Advanced Analytics and Machine Learning

**Getting Started/Training/Tutorials**
- Link to OAA/Oracle Data Miner Workflow GUI Online (free) Tutorial Series on OTN
- Link to OAA/Oracle R Enterprise (free) Tutorial Series on OTN
- Link to Try the Oracle Cloud Now!
- Link to Getting Started w/ ODM blog entry
- Link to New OAA/Oracle Data Mining 2-Day Instructor Led Oracle University course.
- Oracle Data Mining Sample Code Examples

**Additional Resources, Documentation & OTN Discussion Forums**
- Oracle Advanced Analytics Option on OTN page
- OAA/Oracle Data Mining on OTN page, ODM Documentation & ODM Blog
- OAA/Oracle R Enterprise page on OTN page, ORE Documentation & ORE Blog
- Oracle SQL based Basic Statistical functions on OTN
- Oracle R Advanced Analytics for Hadoop (ORAAH) on OTN

**Analytics and Data Summit**, All Analytics, All Data, No Nonsense.
March 12-14, 2019, Redwood Shores, CA
SAVE THE DATE FOR ANALYTICS AND DATA SUMMIT 2019

All Analytics. All Data. No Nonsense.
March 12-14, 2019
www.analyticsanddatasummit.org