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

**Elastic**
- 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.
Architecture for Modern Cloud Data Warehousing

Autonomous Data Warehouse Cloud

Built-in Access Tools

Oracle Machine Learning

Oracle Object Storage Cloud
Flat Files and Staging

3rd Party DI on Oracle Cloud Compute

3rd Party DI On-premises

Oracle Data Integration Platform Cloud

Data Integration Services

Oracle SQL Developer

Developer Tools

Service Console

Service Management
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
Picking a Fast Car ML Analysis

```
select * from CARS;
```

All fields:
- MPG
- CYLINDERS
- DISPLACEMENT
- HORSEPOWER
- WEIGHT
- ACCELERATION
- MODEL YEAR
- ORIGIN

Graph showing relationship between MPG and other variables.
SQL Statistical Functions Examples Not...
Affinity_Card

```sql
-- Run the EXPLAIN routine to get attribute importance results
BEGIN
  DBMS_PREDICTIVE_ANALYTICS.EXPLAIN:
    data_table_name  => 'Cost_Insurance',
    explain_column_name => 'Buy_Insurance',
    result_table_name  => 'ML_explain_output';
END;
```

PL/SQL procedure successfully completed.

 Took 24 sec. Last updated by dberger at May 24 2017, 2:22:02 PM (undefined)

```sql
SELECT attribute_name, explanatory_value, rank 
FROM ML_explain_output 
ORDER BY rank, attribute_name;
```

[Bar chart showing attribute importance]
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
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

“Analytics Consumers”
BI Analysts, Managers
Functional Users (HCM, CRM)

Oracle’s Big Data Cloud Service
“Oracle Machine Learning” Big Data Cloud

Oracle’s Database Cloud
“Oracle Machine Learning” Database Edition

Data Management + Advanced Analytical Platform
Big Data SQL

Oracle’s Machine Learning/Advanced Analytics Platforms

Oracle’s Machine Learning Algorithms Embedded in the Data Management Platforms

“Analytics Producers”
Data Scientists, R Users, Citizen Data Scientists

“Analytics Consumers”
BI Analysts, Managers
Functional Users (HCM, CRM)

Oracle’s Big Data Cloud Service
“Oracle Machine Learning” Big Data Cloud

Oracle’s Database Cloud
“Oracle Machine Learning” Database Edition

Data Management + Advanced Analytical Platform
Big Data SQL

Copyright © 2018, Oracle and/or its affiliates. All rights reserved. |
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

Oracle’s Machine Learning/Advanced Analytics Platforms

Oracle Data Management + Advanced Analytical Platform
Big Data SQL

Oracle Database Cloud
“Oracle Machine Learning” Database Edition
Machine Learning Algorithms,
Statistical Functions + R Integration
for Scalable, Parallel, Distributed, in-DB Execution

Oracle Big Data Cloud Service
“Oracle Machine Learning” Big Data Cloud
ORAAH—Machine Learning Algorithms
Statistical Functions + R Integration
for Scalable, Parallel, Distributed Execution

Copyright © 2018, Oracle and/or its affiliates. All rights reserved.
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 User Profiles

Oracle’s Machine Learning/Advanced Analytics

DBAs

Application Developers

R User, “Data Scientist”

Data Analyst, “Citizen Data Scientist”
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

“Engineered Features”
Derived attributes that reflect domain knowledge—key to best models e.g:
- Counts
- Totals
- Changes over time

Object Store
NoSQL
ORACLE
kafka

Copyright © 2016, Oracle and/or its affiliates. All rights reserved.
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>
<tr>
<td>GLM Regression</td>
<td>640</td>
<td>55s / 512</td>
<td>93s / 144</td>
</tr>
<tr>
<td>Support Vector Machine (IPM solver)</td>
<td>640</td>
<td>404s / 512</td>
<td>1411s / 144</td>
</tr>
<tr>
<td>Support Vector Machine (SGD solver)</td>
<td>640</td>
<td>84s / 256</td>
<td>188s / 72</td>
</tr>
</tbody>
</table>

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 OK 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;
begindbms_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:

```
PREDICTION_PROB... 0.8382936507936...
```
### Build Predictive Models on an Attribute

**Oracle’s Machine Learning Accelerates New Possibilities**

**Machine Learning Model** → **Function**(X₁, X₂, ..., X) → **Y** (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</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>0.000063525000</td>
<td>2.0000</td>
<td>Cashier</td>
<td>VERY HIGH</td>
<td>91</td>
<td></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>0.000058576000</td>
<td>1.0000</td>
<td>Sales</td>
<td>MEDIUM</td>
<td>77</td>
<td></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>6,300.000065225000</td>
<td>2.0000</td>
<td>Administrator</td>
<td>HIGH</td>
<td>64</td>
<td></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>25.0000</td>
<td>0.000068721000</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>78</td>
<td></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>0.000057948000</td>
<td>2.0000</td>
<td>Prof.</td>
<td>HIGH</td>
<td>52</td>
<td></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>999.0000</td>
<td>20,10000060818000</td>
<td>1.0000</td>
<td>PROFESSION</td>
<td>MEDIUM</td>
<td>93</td>
<td></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>8,200.000061322000</td>
<td>1.0000</td>
<td>Caring</td>
<td>HIGH</td>
<td>83</td>
<td></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>2,000.000064742000</td>
<td>1.0000</td>
<td>Prof.</td>
<td>HIGH</td>
<td>54</td>
<td></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>0.000062332000</td>
<td>1.0000</td>
<td>Waiter/Waitress</td>
<td>HIGH</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>CU7127</td>
<td>50.0000</td>
<td>M</td>
<td>WIDOWED</td>
<td>2.0000</td>
<td>29,106.0000</td>
<td>25.0000</td>
<td>0.000065724000</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>68</td>
<td></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>1,100.000659399000</td>
<td>1.0000</td>
<td>Programmer/...</td>
<td>MEDIUM</td>
<td>55</td>
<td></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>0.000066297000</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>78</td>
<td></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>1,186.000</td>
<td>2,600.000669319000</td>
<td>1.0000</td>
<td>School Teacher</td>
<td>HIGH</td>
<td>93</td>
<td></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>0.000066659000</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>98</td>
<td></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>0.000061301000</td>
<td>1.0000</td>
<td>Fireman</td>
<td>MEDIUM</td>
<td>89</td>
<td></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>0.000063381000</td>
<td>1.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>90</td>
<td></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>0.000066803000</td>
<td>1.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>76</td>
<td></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>25.0000</td>
<td>0.000060859000</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>92</td>
<td></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>0.000067749000</td>
<td>1.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>65</td>
<td></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>2,150.00067410000</td>
<td>1.0000</td>
<td>Veternarian</td>
<td>HIGH</td>
<td>93</td>
<td></td>
</tr>
</tbody>
</table>
### Build Predictive Models on an Attribute

**Oracle’s Machine Learning Accelerates New Possibilities**

Machine Learning Model ➔ Function($X_1$, $X_2$, ..., $X$) ➔ $Y$ (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</th>
<th>SALARY</th>
<th>HOUSE_ownership</th>
<th>PROFESSION</th>
<th>LTV_BIN</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>CU8671</td>
<td>46</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>34,040.0000</td>
<td>25.0000</td>
<td>0.000000E+00</td>
<td>63,750...</td>
<td>2.0000</td>
<td>Cashier</td>
<td>VERY HIGH</td>
<td>91</td>
</tr>
<tr>
<td>CU7115</td>
<td>27</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>20,843.2500</td>
<td>25.0000</td>
<td>0.000000E+00</td>
<td>58,750...</td>
<td>1.0000</td>
<td>PROF-14</td>
<td>MEDIUM</td>
<td>77</td>
</tr>
<tr>
<td>CU7117</td>
<td>30</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>28,306.5000</td>
<td>349.0000</td>
<td>0.000000E+00</td>
<td>65,225...</td>
<td>2.0000</td>
<td>Administrator</td>
<td>HIGH</td>
<td>64</td>
</tr>
<tr>
<td>CU7118</td>
<td>33</td>
<td>M</td>
<td>DIVORCED</td>
<td>1.0000</td>
<td>26,480.2500</td>
<td>25.0000</td>
<td>0.000000E+00</td>
<td>66,721...</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>78</td>
</tr>
<tr>
<td>CU7120</td>
<td>21</td>
<td>M</td>
<td>DIVORCED</td>
<td>3.0000</td>
<td>22,012.0000</td>
<td>25.0000</td>
<td>0.000000E+00</td>
<td>57,648...</td>
<td>1.0000</td>
<td>Professor</td>
<td>HIGH</td>
<td>52</td>
</tr>
<tr>
<td>CU7121</td>
<td>32</td>
<td>F</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>20,904.5000</td>
<td>999.0000</td>
<td>0.000000E+00</td>
<td>60,818...</td>
<td>1.0000</td>
<td>PROF-38</td>
<td>MEDIUM</td>
<td>93</td>
</tr>
<tr>
<td>CU7123</td>
<td>35</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>22,330.5000</td>
<td>92.0000</td>
<td>0.000000E+00</td>
<td>61,322...</td>
<td>1.0000</td>
<td>Clinics</td>
<td>HIGH</td>
<td>83</td>
</tr>
<tr>
<td>CU7124</td>
<td>44</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>23,085.5000</td>
<td>25.0000</td>
<td>0.000000E+00</td>
<td>60,742...</td>
<td>1.0000</td>
<td>PROF-1</td>
<td>HIGH</td>
<td>54</td>
</tr>
<tr>
<td>CU7125</td>
<td>63</td>
<td>M</td>
<td>MARRIED</td>
<td>1.0000</td>
<td>25,383.0000</td>
<td>25.0000</td>
<td>0.000000E+00</td>
<td>62,332...</td>
<td>1.0000</td>
<td>Waiter/Waitress</td>
<td>HIGH</td>
<td>74</td>
</tr>
<tr>
<td>CU7127</td>
<td>58</td>
<td>M</td>
<td>WIDOWED</td>
<td>2.0000</td>
<td>29,106.5000</td>
<td>25.0000</td>
<td>0.000000E+00</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</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>15,084.7500</td>
<td>708.0000</td>
<td>1.000000E+00</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</td>
<td>M</td>
<td>MARRIED</td>
<td>0.0000</td>
<td>30,251.7500</td>
<td>25.0000</td>
<td>0.000000E+00</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</td>
<td>F</td>
<td>DIVORCED</td>
<td>4.0000</td>
<td>25,222.7500</td>
<td>1,186.0000</td>
<td>0.000000E+00</td>
<td>66,931...</td>
<td>1.0000</td>
<td>School Teacher</td>
<td>HIGH</td>
<td>93</td>
</tr>
<tr>
<td>CU7132</td>
<td>39</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>32,064.7500</td>
<td>25.0000</td>
<td>0.000000E+00</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</td>
<td>F</td>
<td>MARRIED</td>
<td>3.0000</td>
<td>15,425.2500</td>
<td>25.0000</td>
<td>0.000000E+00</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</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>20,695.2500</td>
<td>25.0000</td>
<td>0.000000E+00</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</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>21,009.7500</td>
<td>25.0000</td>
<td>0.000000E+00</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</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>22,064.7500</td>
<td>25.0000</td>
<td>0.000000E+00</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</td>
<td>M</td>
<td>SINGLE</td>
<td>0.0000</td>
<td>20,237.2500</td>
<td>25.0000</td>
<td>0.000000E+00</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</td>
<td>F</td>
<td>MARRIED</td>
<td>5.0000</td>
<td>24,152.5000</td>
<td>25.0000</td>
<td>2,150.0000E+00</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** ➔ **Function**($X_1, X_2, \ldots, X$) ➔ **Y2** (BankFunds) ➔ **Y** (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</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>1</td>
<td>46.000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>34,040.000</td>
<td>25.000</td>
<td>0.0000</td>
<td>4,500,000</td>
<td>63,760</td>
<td>2.0000</td>
<td>Cashier</td>
<td>VERY HIGH</td>
<td>91</td>
</tr>
<tr>
<td>2</td>
<td>27.000</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>1,500</td>
<td>58,573</td>
<td>1.0000</td>
<td>Prof-14</td>
<td>MEDIUM</td>
<td>77</td>
</tr>
<tr>
<td>3</td>
<td>20.000</td>
<td>F</td>
<td>DIVORCED</td>
<td>5.0000</td>
<td>28,306.5000</td>
<td>349.000</td>
<td>6,300.000</td>
<td>35,000,000</td>
<td>65,226</td>
<td>2.0000</td>
<td>Administrator</td>
<td>HIGH</td>
<td>64</td>
</tr>
<tr>
<td>4</td>
<td>33.000</td>
<td>M</td>
<td>DIVORCED</td>
<td>1.0000</td>
<td>20,460.000</td>
<td>25.000</td>
<td>0.0000</td>
<td>150,000</td>
<td>66,721</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>78</td>
</tr>
<tr>
<td>5</td>
<td>21.000</td>
<td>M</td>
<td>DIVORCED</td>
<td>3.0000</td>
<td>22,012.000</td>
<td>25.000</td>
<td>0.0000</td>
<td>-5,000</td>
<td>57,548</td>
<td>1.0000</td>
<td>Professor</td>
<td>HIGH</td>
<td>52</td>
</tr>
<tr>
<td>6</td>
<td>32.000</td>
<td>M</td>
<td>MARRIED</td>
<td>2.0000</td>
<td>20,904.5000</td>
<td>999.000</td>
<td>20,100.000</td>
<td>210,000</td>
<td>60,818</td>
<td>1.0000</td>
<td>PROF-38</td>
<td>MEDIUM</td>
<td>93</td>
</tr>
<tr>
<td>7</td>
<td>35.000</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>8,500,000</td>
<td>61,322</td>
<td>1.0000</td>
<td>Clerico</td>
<td>HIGH</td>
<td>83</td>
</tr>
<tr>
<td>8</td>
<td>44.000</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>2,500,000</td>
<td>60,742</td>
<td>1.0000</td>
<td>PROF-1</td>
<td>HIGH</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>63.000</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>25,000</td>
<td>62,332</td>
<td>1.0000</td>
<td>Waiter/Waitress</td>
<td>HIGH</td>
<td>74</td>
</tr>
<tr>
<td>10</td>
<td>58.000</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>11,000</td>
<td>65,224</td>
<td>1.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>68</td>
</tr>
<tr>
<td>11</td>
<td>26.000</td>
<td>M</td>
<td>SINGLE</td>
<td>1.0000</td>
<td>15,084.7500</td>
<td>708.000</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>12</td>
<td>28.000</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>17,000</td>
<td>66,207</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>78</td>
</tr>
<tr>
<td>13</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>14</td>
<td>39.000</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>0</td>
<td>66,659</td>
<td>1.0000</td>
<td>Nurse</td>
<td>VERY HIGH</td>
<td>98</td>
</tr>
<tr>
<td>15</td>
<td>26.000</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>-4,500</td>
<td>61,301</td>
<td>1.0000</td>
<td>Fireman</td>
<td>MEDIUM</td>
<td>89</td>
</tr>
<tr>
<td>16</td>
<td>38.000</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>9,000</td>
<td>63,381</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>90</td>
</tr>
<tr>
<td>17</td>
<td>42.000</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>6,500</td>
<td>66,803</td>
<td>0.0000</td>
<td>Nurse</td>
<td>MEDIUM</td>
<td>76</td>
</tr>
<tr>
<td>18</td>
<td>36.000</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>12,000</td>
<td>69,859</td>
<td>0.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>92</td>
</tr>
<tr>
<td>19</td>
<td>23.000</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>7,500</td>
<td>67,749</td>
<td>0.0000</td>
<td>Nurse</td>
<td>HIGH</td>
<td>65</td>
</tr>
<tr>
<td>20</td>
<td>38.000</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>2,200,000</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**

<table>
<thead>
<tr>
<th>Machine Learning Models</th>
<th>Function($X_1, X_2, \ldots, X$)</th>
<th>$Y_2$ (BankFunds)</th>
<th>$Y$ (LTV BIN); Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>CUST_ID</td>
<td>AGE</td>
<td>SEX</td>
<td>MARITAL_STATUS</td>
</tr>
<tr>
<td>1</td>
<td>CU8817</td>
<td>46.0000</td>
<td>M</td>
</tr>
<tr>
<td>2</td>
<td>CU7115</td>
<td>27.0000</td>
<td>M</td>
</tr>
<tr>
<td>3</td>
<td>CU7117</td>
<td>30.0000</td>
<td>F</td>
</tr>
<tr>
<td>4</td>
<td>CU7118</td>
<td>33.0000</td>
<td>M</td>
</tr>
<tr>
<td>5</td>
<td>CU7120</td>
<td>21.0000</td>
<td>M</td>
</tr>
<tr>
<td>6</td>
<td>CU7121</td>
<td>32.0000</td>
<td>F</td>
</tr>
<tr>
<td>7</td>
<td>CU7123</td>
<td>35.0000</td>
<td>M</td>
</tr>
<tr>
<td>8</td>
<td>CU7124</td>
<td>44.0000</td>
<td>F</td>
</tr>
<tr>
<td>9</td>
<td>CU7125</td>
<td>55.0000</td>
<td>M</td>
</tr>
<tr>
<td>10</td>
<td>CU7127</td>
<td>50.0000</td>
<td>M</td>
</tr>
<tr>
<td>11</td>
<td>CU7128</td>
<td>26.0000</td>
<td>M</td>
</tr>
<tr>
<td>12</td>
<td>CU7129</td>
<td>28.0000</td>
<td>M</td>
</tr>
<tr>
<td>13</td>
<td>CU7130</td>
<td>47.0000</td>
<td>F</td>
</tr>
<tr>
<td>14</td>
<td>CU7132</td>
<td>39.0000</td>
<td>M</td>
</tr>
<tr>
<td>15</td>
<td>CU7133</td>
<td>26.0000</td>
<td>F</td>
</tr>
<tr>
<td>16</td>
<td>CU7134</td>
<td>38.0000</td>
<td>M</td>
</tr>
<tr>
<td>17</td>
<td>CU7135</td>
<td>42.0000</td>
<td>M</td>
</tr>
<tr>
<td>18</td>
<td>CU7136</td>
<td>36.0000</td>
<td>M</td>
</tr>
<tr>
<td>19</td>
<td>CU7137</td>
<td>23.0000</td>
<td>M</td>
</tr>
<tr>
<td>20</td>
<td>CU7138</td>
<td>38.0000</td>
<td>F</td>
</tr>
</tbody>
</table>
The Core Ingredients of Good Machine Learning

Domain Knowledge + Data

Machine Learning Algorithms

Insights, Predictions
Most Important Factor in Machine Learning? **Deployment!!**

A “Thinking” Database
Community and Resources
ANALYTICS AND DATA SUMMIT 2018

All Analytics. All Data. No Nonsense.
March 20-22, 2018
www.analyticstechsummit.org

We’ve changed our name! Formerly called the BIWA Summit with the Spatial and Graph Summit.
Same great technical content – great new name!
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** preso
- 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**

**Oracle Help Center 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 20-22, 2018, Redwood Shores, CA