

Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics/ Oracle Data Miner Hands on Lab

Charlie Berger, MS Engineering, MBA, Sr. Director Product Management, Advanced Analytics and Machine Learning

charlie.berger@oracle.com

www.twitter.com/CharlieDataMine

Tim Vlamis, Consultant, Vlamis Software Solutions, Inc.

Karl Rexer, President, Rexer Analytics



Accelerate Your
Digital Transformation
in the Cloud



Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab



- **Lessons—Novices**

1. OAA Quick Overview
2. Data Mining Concepts Briefly
3. Quick Oracle Data Miner GUI Demo
4. Take off! Do as many Tutorials as you can in the 2 hours HOL
5. Ask questions! We're all here to help and discuss use cases!



- **Take off!—Intermediate/Experts**

1. Environment

- Oracle 12c on the **Oracle Database Cloud**
- Will be using SQL Developer 4.2 EA

2. Do 3-5 Tutorials

- Instructors will walk around helping

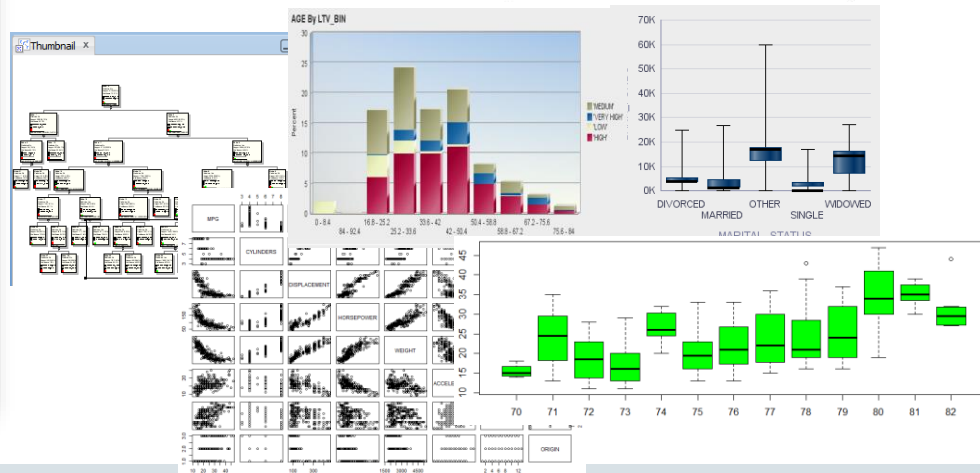
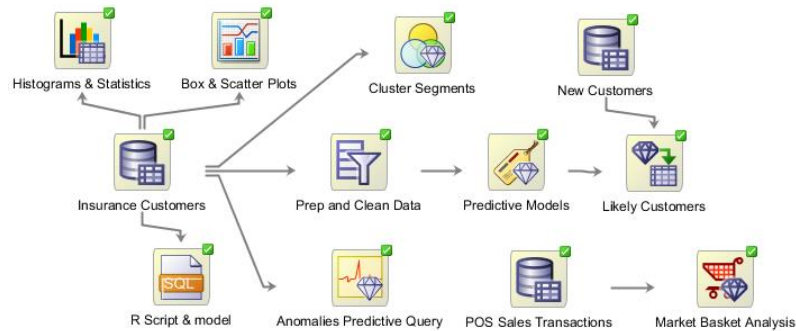
Oracle's Advanced Analytics

Fastest Way to Deliver Scalable Enterprise-wide Predictive Analytics



Key Features

- Parallel, scalable data mining 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



Google “Oracle Advanced Analytics”

Welcome Charlie

Account Sign Out Help Country Communities I am a... I want to... Search

Products Solutions Downloads Store Support Training Partners About OTN

Oracle Technology Network > Database > Options > Advanced Analytics > Overview

- Database 12c
- Database In-Memory
- Multitenant
- Options
- Application Development
- Big Data Appliance
- Data Warehousing & Big Data
- Database Appliance
- Database Cloud
- Exadata Database Machine
- High Availability
- Manageability
- Migrations
- Security
- Unstructured Data
- Upgrades
- Windows
- Database Technology Index

Overview Downloads Documentation Community Learn More

Oracle Advanced Analytics

Scalable enterprise-wide predictive analytics

Architecture Overview

Oracle Advanced Analytics 12c delivers parallelized in-database implementations of data mining algorithms and integration with open source R. Data analysts use Oracle Data Miner GUI and R to build and evaluate predictive models and leverage R packages and graphs. Application developers deploy Oracle Advanced Analytics models using SQL data mining functions and R. With the Oracle Advanced Analytics option, Oracle extends the Oracle Database to an *scalable analytical platform* that mines more data and data types, eliminates data movement, and preserves security to anticipate customer behavior, detect patterns, and deliver actionable insights. Oracle Big Data SQL adds new big data sources and Oracle R Advanced Analytics for Hadoop provides algorithms that run on Hadoop.

Oracle Advanced Analytics, a combination of **Oracle Data Mining** and **Oracle R Enterprise**, delivers predictive analytics, data mining, text mining, statistical analysis, advanced visualization capabilities, and interactive graphics inside the database. It provides a scalable analytical platform that mines more data and data types, eliminates data movement, and preserves security, and savings. Data analysts, data scientists, and database administrators can develop and deploy predictive models and methodologies inside the database and gain competitive advantage.

VISIT ORACLE AT COLLABORATE 16

APRIL 10-14, 2016
MANDALAY BAY RESORT & CASINO
LAS VEGAS, NV

REGISTER NOW

Oracle Database Cloud

Get Started >

Get the Latest Oracle Database 12c Tutorials

Plug into the Cloud

Access Now >

Oracle Data Mining

Scalable in-database predictive analytics



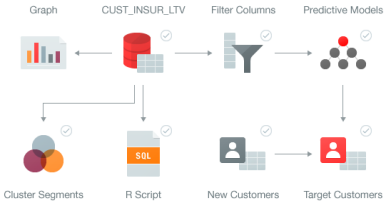
Oracle R Advanced Analytics for Hadoop

Overview

Oracle Data Mining (ODM), a component of the **Oracle Advanced Analytics Database Option**, provides powerful data mining algorithms that enable analysts to discover insights, make predictions and leverage their Oracle investment. With ODM, you can build and apply predictive models inside Database to help you predict customer behavior, target your best custom develop customer profiles, identify cross-selling opportunities and detect and prevent fraud.

Oracle Data Miner

Overview Downloads Documentation Community Learn More




Oracle R Enterprise

Overview Downloads Documentation Community Learn More

News Fact: Oracle R Enterprise 1.5 is now generally available!

News: Automated Trading Strategies using R

Oracle R Enterprise, a component of the **Oracle Advanced Analytics Option**, makes the open source R statistical programming language and environment ready for the enterprise and big data. Designed for problems involving large volumes of data, Oracle R Enterprise integrates R with Oracle Database. R users can run R commands and scripts for statistical and graphical analyses on data stored in Oracle Database. R users can develop, refine, and deploy R scripts that leverage the parallelism and scalability of Oracle Database to automate data analysis. Data analysts and data scientists can run R packages and develop and operationalize R scripts for analytical applications in one step—without having to learn SQL. Oracle R Enterprise performs function pushdown for in-database execution of base R and popular R package functions. Being integrated with Oracle Database, Oracle R Enterprise can run any R package via embedded R while the database manages the data served to the R engines.

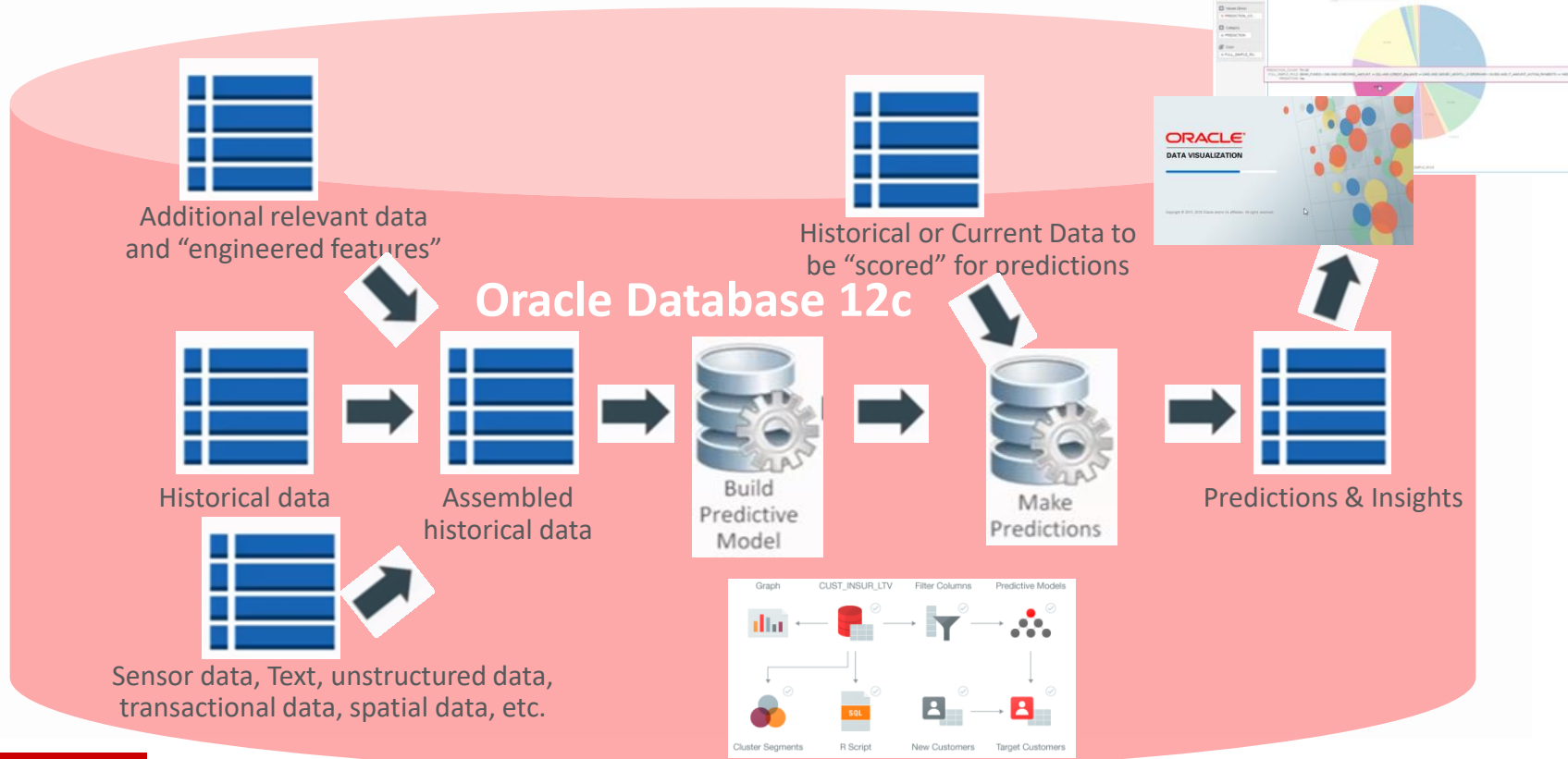


Advanced Analytics



Oracle Data Mining/ Machine Learning/Predictive Analytics

Data Preparation & Adv. Analytical Process Runs In-Database



Oracle Advanced Analytics 12.2

Model Build Time Performance

Prelim/Unofficial



NEW IN
12.2

OAA 12.2 Algorithms

	<u>Rows</u> (Ms)	<u>T7-4</u> (Sparc & Solaris) <u>Model Build Time</u> (Secs / Degree of Parallelism)	<u>X5-4</u> (Intel and Linux) <u>Model Build Time</u> (Secs / Degree of Parallelism)
Attributes Importance	640	28s / 512	44s / 72
K Means Clustering	640	161s / 256	268s / 144
Expectation Maximization	159	455s / 512	588s / 144
Naive Bayes Classification	320	17s / 256	23s / 72
GLM Classification	640	154s / 512	363s / 144
GLM Regression	640	55s / 512	93s / 144
Support Vector Machine (IPM solver)	640	404s / 512	1411s / 144
Support Vector Machine (SGD solver)	640	84s / 256	188s / 72

Wow! That's Fast!

ORACLE

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 9K 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.

Oracle's Advanced Analytics and Machine Learning Platform

Multiple interfaces across platforms — SQL, R, GUI, Dashboards, Apps

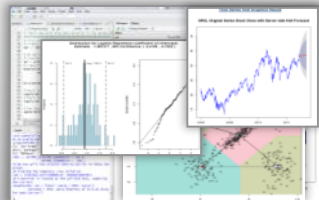
Users



Information Producers

R programmers

R Client



Data & Business Analysts

SQLDEV/
Oracle Data Miner

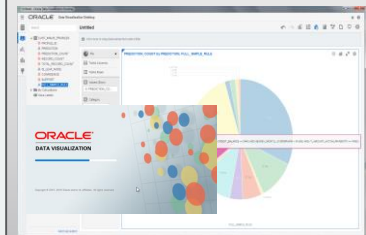


Information Consumers

Business Analysts/Mgrs

Domain End Users (HCM, CRM)

OBIEE/DV



Applications



Platform

Hadoop

ORAAH
Parallel,
distributed
algorithms

Oracle Database Enterprise Edition

Oracle Advanced Analytics - Database Option
SQL Data Mining, ML & Analytic Functions + R Integration
for Scalable, Distributed, Parallel in-DB ML Execution

Oracle Cloud



Oracle Database 12c



ORACLE®

Oracle Data Miner GUI

Easy to Use “Citizen Data Scientist”



Advanced Analytics



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

The screenshot displays the Oracle Data Miner interface within the SQL Developer environment. The main workspace shows a workflow diagram for 'BUY_INSURANCE'. The workflow starts with 'Explore Data' on 'CUST_INSR_LTV1', followed by 'Filter Columns', 'Multiple Classification Models', and 'Most Likely Customers'. The 'Script Output' window shows a SQL query and its results:

```
SQL
select * from
(select POLICYNUMBER, round(prob_fraud*100,2) percent_fraud,
1 POLICYNUMBER | PERCENT_FRAUD | Rnk
1 654 61.87 1
2 11068 57.37 2
3 7435 55.47 3
```

The 'Rule' window shows a rule definition for 'Most Likely Customers' based on 'BANK_FUNDS' and 'CHECKING_AMOUNT'.

```
Rule: Surrogates - Target Values
Node Rule:
If BANK_FUNDS > 246
And CHECKING_AMOUNT > 282
And MONEY_MONTHLY_OVERDRAWN <= 54.215
Then No
Confidence 0.8515671200473093
Support 0.1832863464217212
```



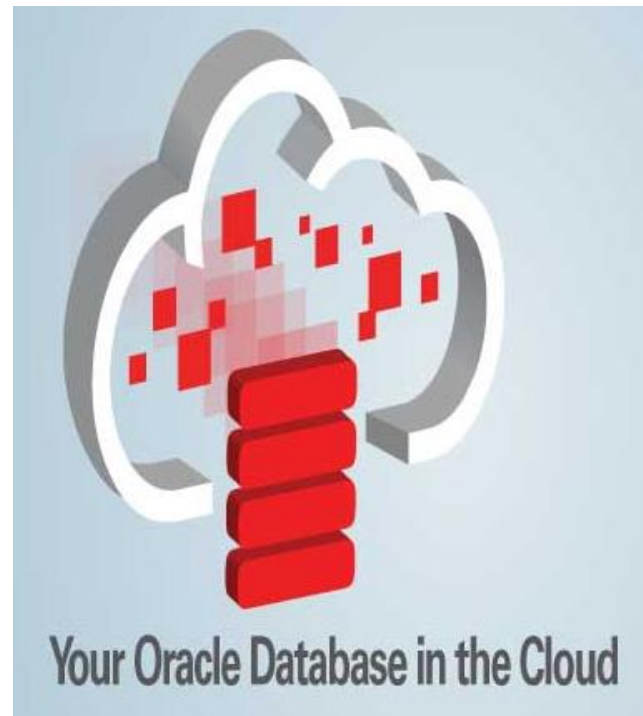
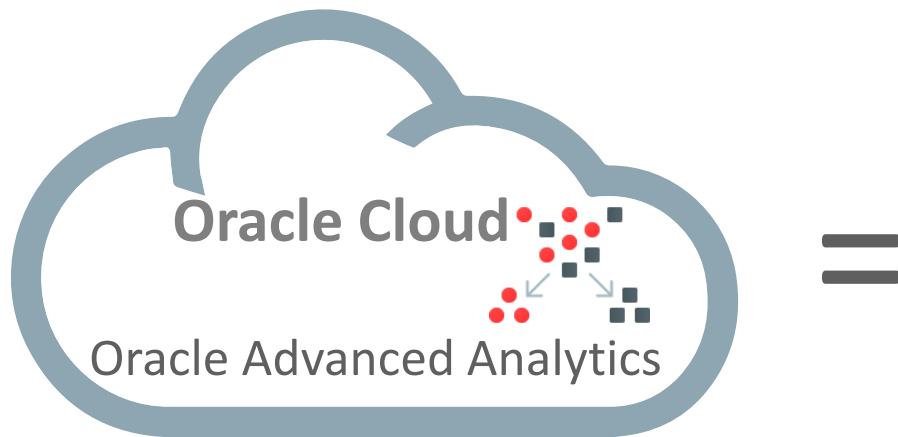

Take off!—Intermediate/Experts

Quick Set up Overview

Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab

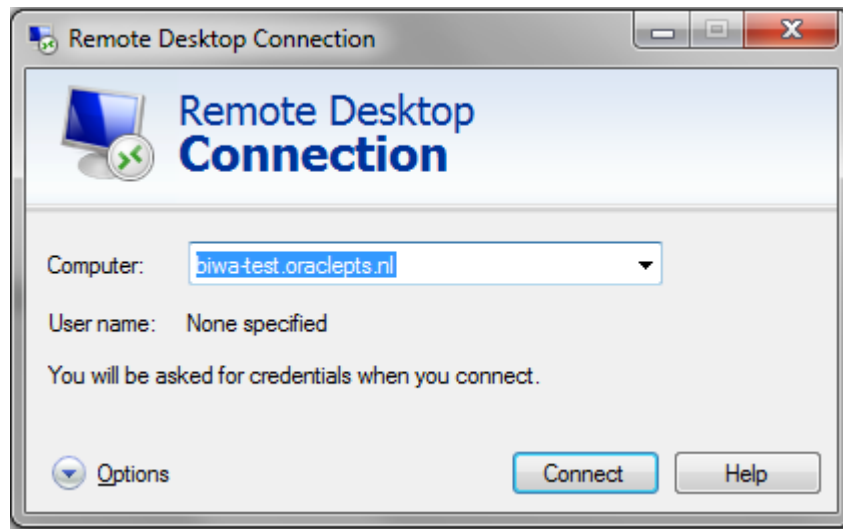
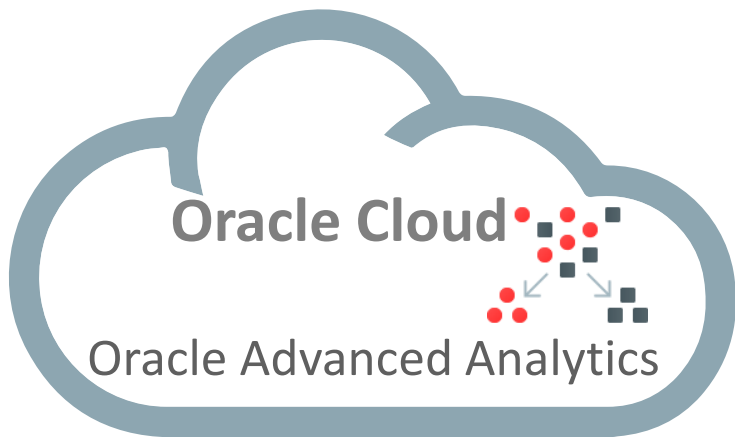
- We're using the Oracle Database Cloud!



Learn Predictive Analytics in 2 Hours!

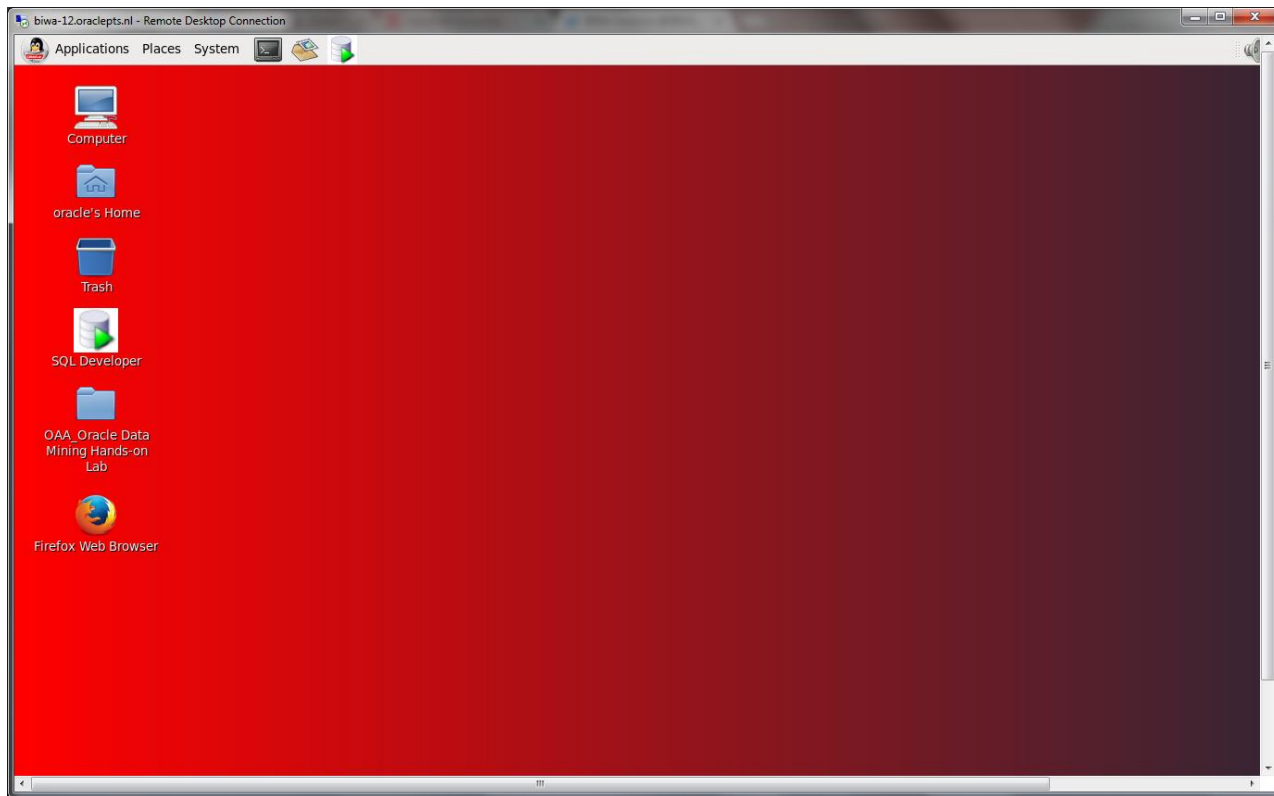
Oracle Advanced Analytics Hands on Lab

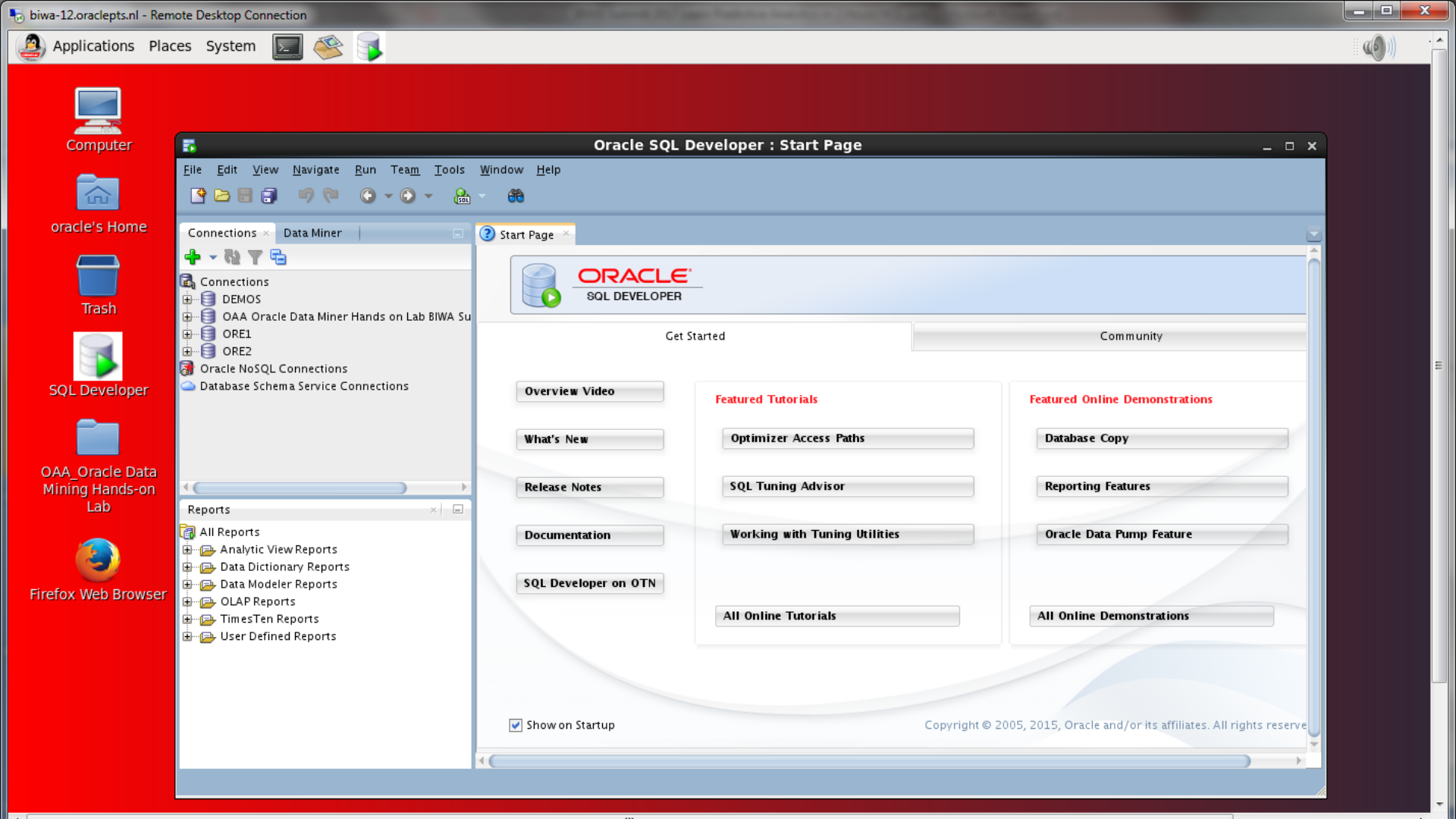
- We're using the Oracle Database Cloud!

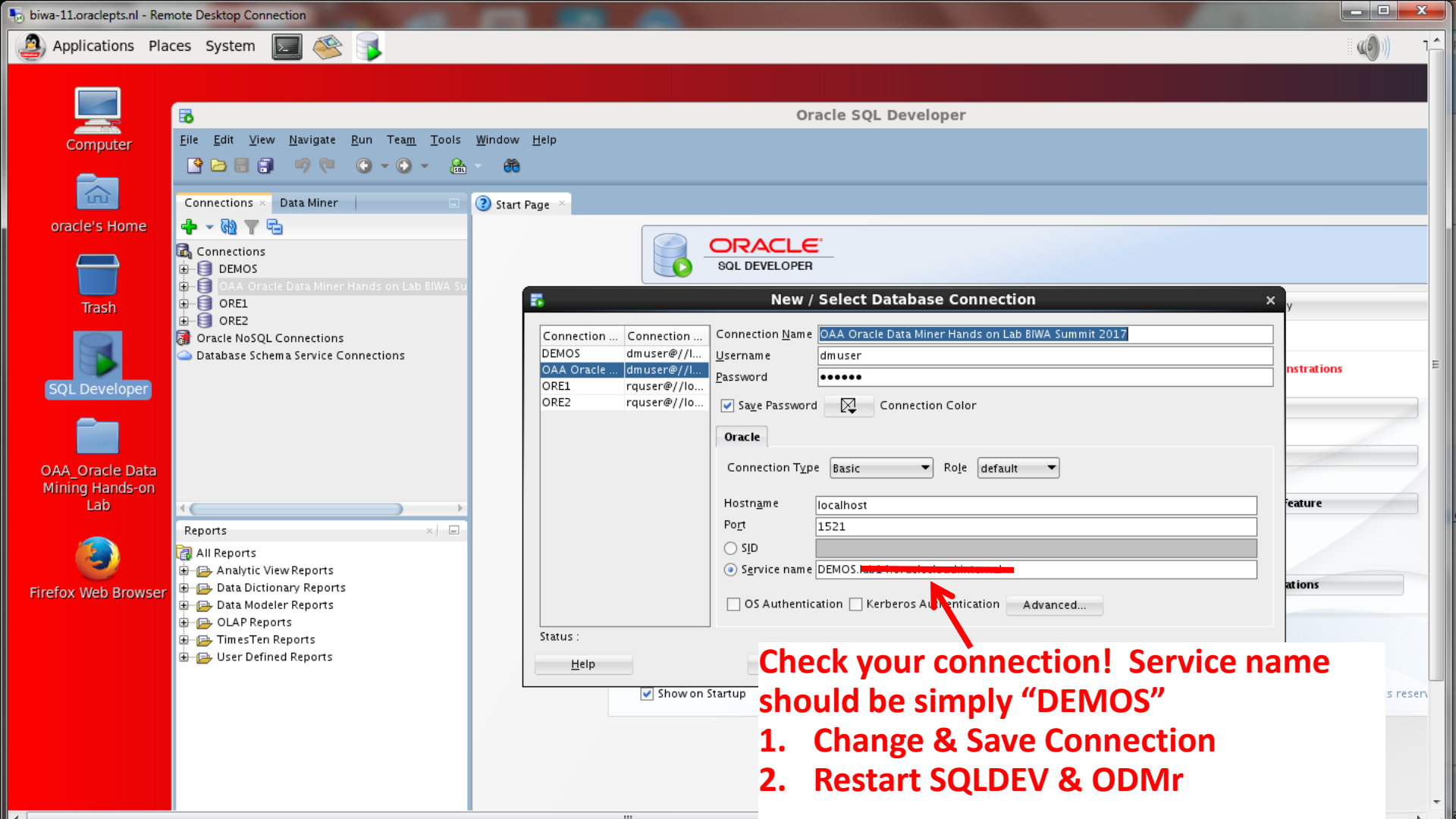


Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab







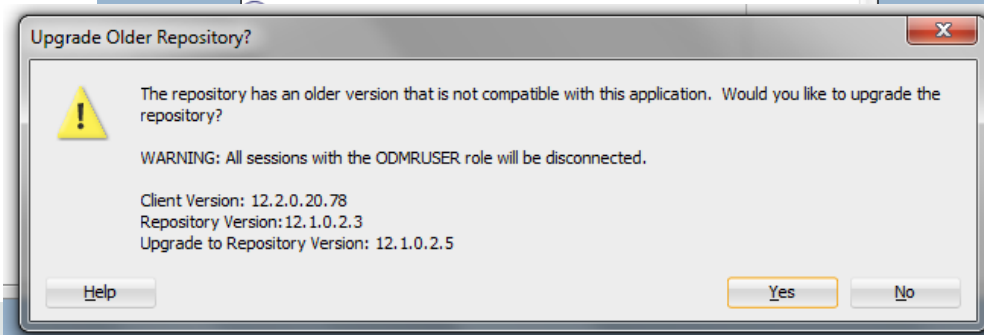
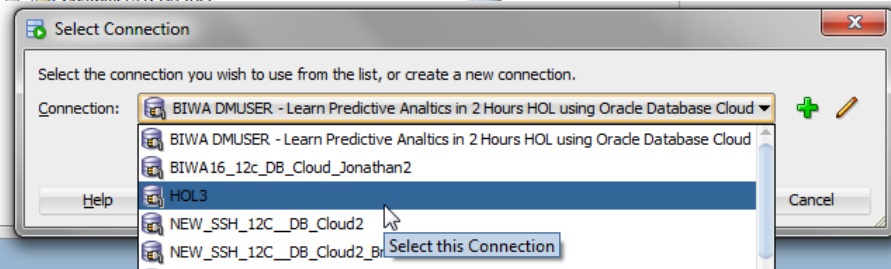
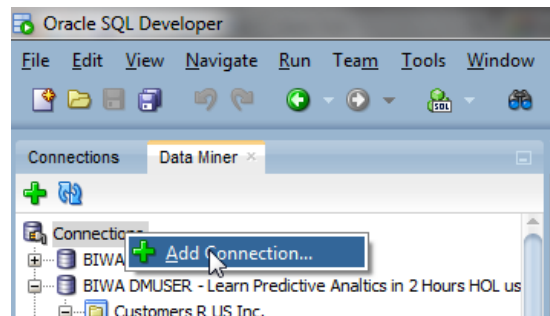
Check your connection! Service name should be simply "DEMOS"

- 1. Change & Save Connection**
- 2. Restart SQLDEV & ODMr**

Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab

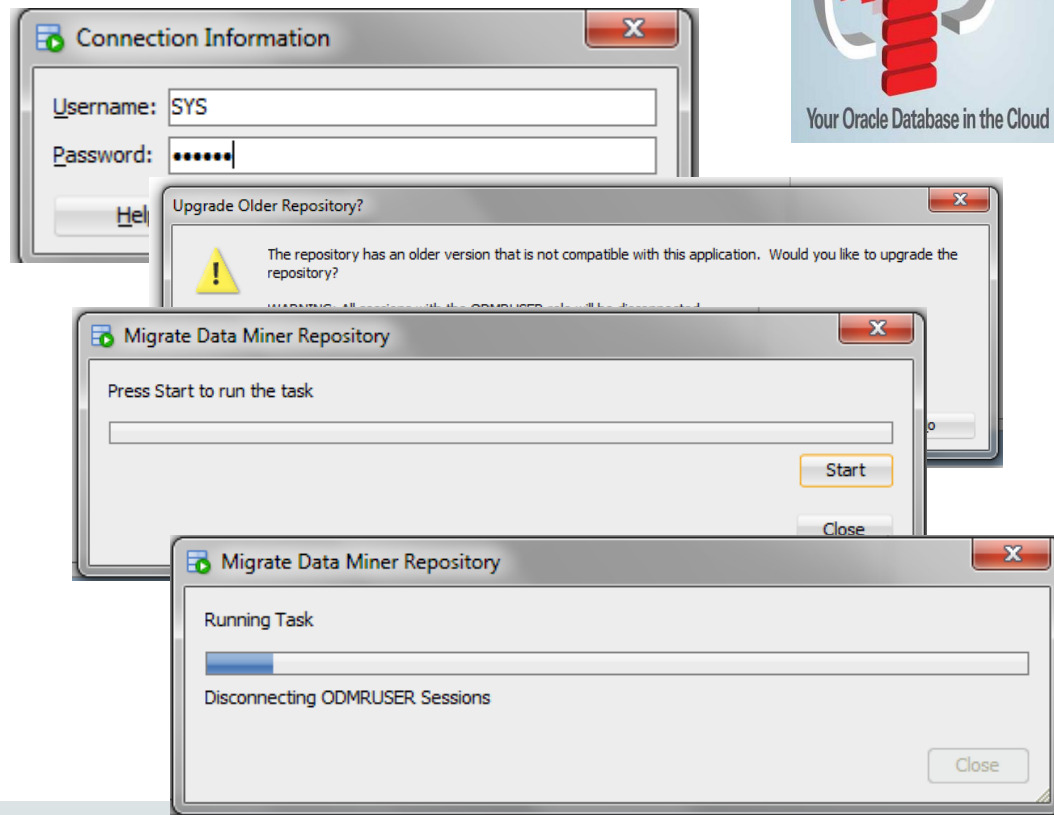
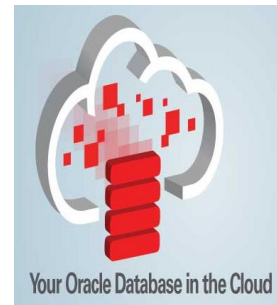
- Step 1—Install SQLDEV 4.1.3
- Step 2—Connect to Oracle Database Cloud
 - 1. Go to Oracle Data Miner & create a NEW Connection e.g. HOL"N"
 - Select HOL"N" from drop down menu
 - Optionally may need to upgrade older Data Mining repository (may take 3 mins)
 - You are done!



Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab

- Step 1—Install SQLDEV 4.1.3
- Step 2—Connect to Oracle Database Cloud
 - Connect as SYS/Welcome#1
 - Start to run task
 - Running task (may take 3 mins)



Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab

- Step 1—Install SQLDEV 4.1.3
- Step 2—Connect to Oracle Database Cloud
- Step 3—Start HOL!

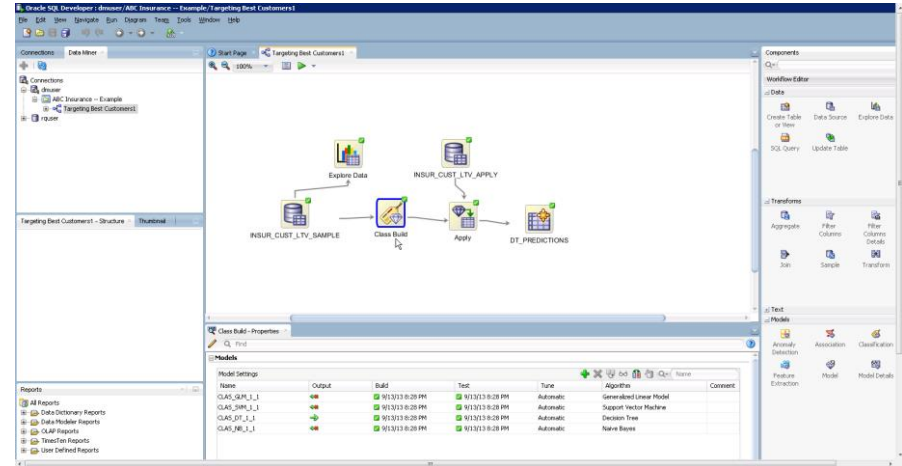
- dmuser/dmuser
- Demo data for learning
- Follow 3-5 OBE Online Tutorials

1. [Using Oracle Data Miner 4.1](#)
2. [Star Schema Mining Using Oracle Data Miner 4.1](#)
3. [Text Mining with an EM Clustering Model Using Data Miner 4.1](#)
4. **Anomaly Detection (CLAIMS)** *See Instructor for assistance*
5. **Market Basket Analysis (SH.SALES)** *See Instructor for assistance*

Cloud Computing
ORACLE
SQL Developer
4.1.3



Exchange Forums **Download** f y d



OAA/Oracle Data Miner 4.2 HOL

We're Using the Oracle by Example Free Online Tutorials

- Google “Oracle Data Miner”
- Scroll down to bottom of page & launch tutorials

– https://apexapps.oracle.com/pls/apex/f?p=44785:24:::NO::P24_CONTENT_ID,P24_PREV_PAGE:11925,2

Demos and Training



2-Day Oracle University Course



Oracle Data Miner Workflow GUI Tutorials

Oracle Data Miner on Big Data Lite VM Tutorial



OAA YouTube demos



Sample Programs

Oracle Data Miner/ SQL Developer Extension



Drag & Drop Workflow for Creating and Deploying in-
Database Analytical Methodologies

Oracle Data Miner GUI

The Oracle Data Miner is an Oracle SQL Developer extension that enables data analysts, business analysts and data scientists to work directly with data inside the database using in-database data mining algorithms of **Oracle Advanced Analytics** and it's **Oracle Data Mining** component SQL functions. Data analyst can use Oracle Data Miner's "drag and drop" workflow and component pallet to explore the data in graphs, build and evaluate data mining models, apply the predictive models to new data and then deploy the OAA/Oracle Data Mining's predictions and insights

OAA/Oracle Data Miner 4.1 HOL

Uses Oracle by Example Free Online Tutorials

- There are 6 Tutorials

- The first tutorial is already done for you

- Recommend doing 3-5 Tutorials

1. [Using Oracle Data Miner 4.1](#)

2. [Star Schema Mining Using Oracle Data Miner 4.1](#)

3. [Text Mining with an EM Clustering Model Using Data Miner 4.1](#)

4. **Anomaly Detection (CLAIMS)** *See Instructor for assistance*

5. **Market Basket Analysis (SH.SALES)** *See Instructor for assistance*

Content List (click links below to view content)

08-SEP-2015
15 mins
★★★★★

[Setting Up Oracle Data Miner 4.1](#)

This tutorial covers the process of setting up Oracle Data Miner for use within Oracle SQL Developer 4.1.

08-SEP-2015
45 mins
★★★★★

[Using Oracle Data Miner 4.1](#)

This tutorial covers the use of Oracle Data Miner 4.1 to perform data mining against Oracle Database 12c. In this lesson, you examine the Oracle Data Miner GUI. The Oracle Data Miner GUI is included as an extension of Oracle SQL Developer, version 4.1.

08-SEP-2015
30 mins
★★★★★

[Star Schema Mining Using Oracle Data Miner 4.1](#)

This tutorial covers the use of Oracle Data Miner 4.1 to perform star schema mining activities against Oracle Database 12c Release 2.

08-SEP-2015
30 mins
★★★★★

[Text Mining with an EM Clustering Model Using Data Miner 4.1](#)

In this lesson, you learn how to use the EM algorithm in a clustering model while leveraging text mining enhancements that are included in Oracle Database 12c Release 2.

08-SEP-2015
30 mins
★★★★★

[Using Logistic Regression Models \(GLM\) to Predict Customer Affinity](#)

This tutorial covers the use of Oracle Data Miner 4.1 to leverage enhancements to the Oracle implementation of Generalized Linear Models (GLM) to predict customer affinity.

08-SEP-2015
30 mins
★★★★★

[Using the SQL Query Node With Oracle Data Miner 4.1](#)

This tutorial covers the use of the new SQL Query Node in an Oracle Data Miner 4.1 workflow.

08-SEP-2015
30 mins
★★★★★

[Using Predictive Queries With Oracle Data Miner 4.1](#)

This tutorial covers the use of Predictive Queries against mining data by Oracle Data Miner 4.1.

08-SEP-2015
30 mins
★★★★★

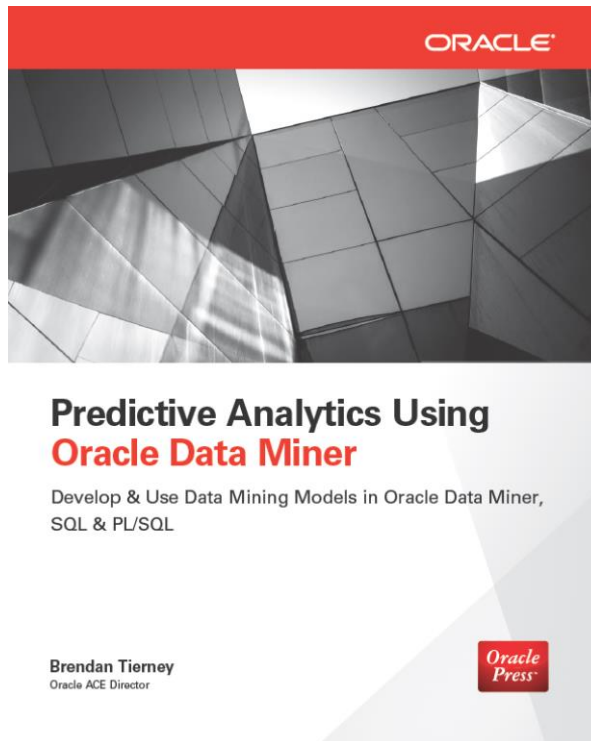
[Mining JSON Data Using Oracle Data Miner 4.1](#)

This tutorial covers the use of the JSON Query Node in an Oracle Data Miner 4.1 workflow in order to mine this Big Data format.

Great book on Oracle Advanced Analytics

Available on Amazon or from Author

- Predictive Analytics Using Oracle Data Miner: Develop for ODM in SQL & PL/SQL



OAA/Oracle Data Miner 4.1 HOL

Setting Up Oracle Data Miner

Setting Up Oracle Data Miner 4.1

Done ✓

The image shows the Oracle Data Miner 4.1 setup wizard interface. On the left, there is a sidebar with 'Options' (Expand All Topics, Hide All Media) and a list of steps: 'Before You Begin', 'Create a Data Miner User Account', 'Create a SQL Developer Connection for the Data Miner User', 'Install the Data Miner Repository', and 'Want to Learn More?'. A red arrow points from the 'Before You Begin' step to the 'New / Select Database Connection' dialog box. The dialog box is titled 'New / Select Database Connection' and contains fields for 'Connection Name' (admin), 'Username' (SYS), and 'Password' (masked). It also has checkboxes for 'Save Password' and 'Connection Color'. Below these, the 'Oracle' tab is selected, showing 'Connection Type' (Basic), 'Role' (SYSDBA), 'Hostname' (localhost), 'Port' (1521), and 'SID' (orcl). There are also checkboxes for 'OS Authentication' and 'Kerberos Authentication', and an 'Advanced...' button. At the bottom, there are buttons for 'Save', 'Clear', 'Test' (highlighted with a yellow border), 'Connect', and 'Cancel'.

OAA/Oracle Data Miner 4.1 HOL

Setting Up Oracle Data Miner

Setting Up Oracle Data Miner 4.1

Done ✓

Options

[Expand All Topics](#)

[Hide All Media](#)

- Before You Begin
- Create a Data Miner User Account
- Create a SQL Developer Connection for the Data Miner User
- Install the Data Miner Repository
- Want to Learn More?

Create User

User | Granted Roles | System Privileges | Quotas | SQL

User Name:

New Password:

Confirm Password:

☐ Password Expired (user must change next login)

☐ Operating System User

☐ Account is Locked

☐ Edition Enabled

Default Tablespace:

Temporary Tablespace:

Connections

[New Connection...](#)

[Import Connections...](#)

[Export Connections...](#)

[Create Local Connections](#)

[Disconnect Connections](#)

Install Data Miner Repository

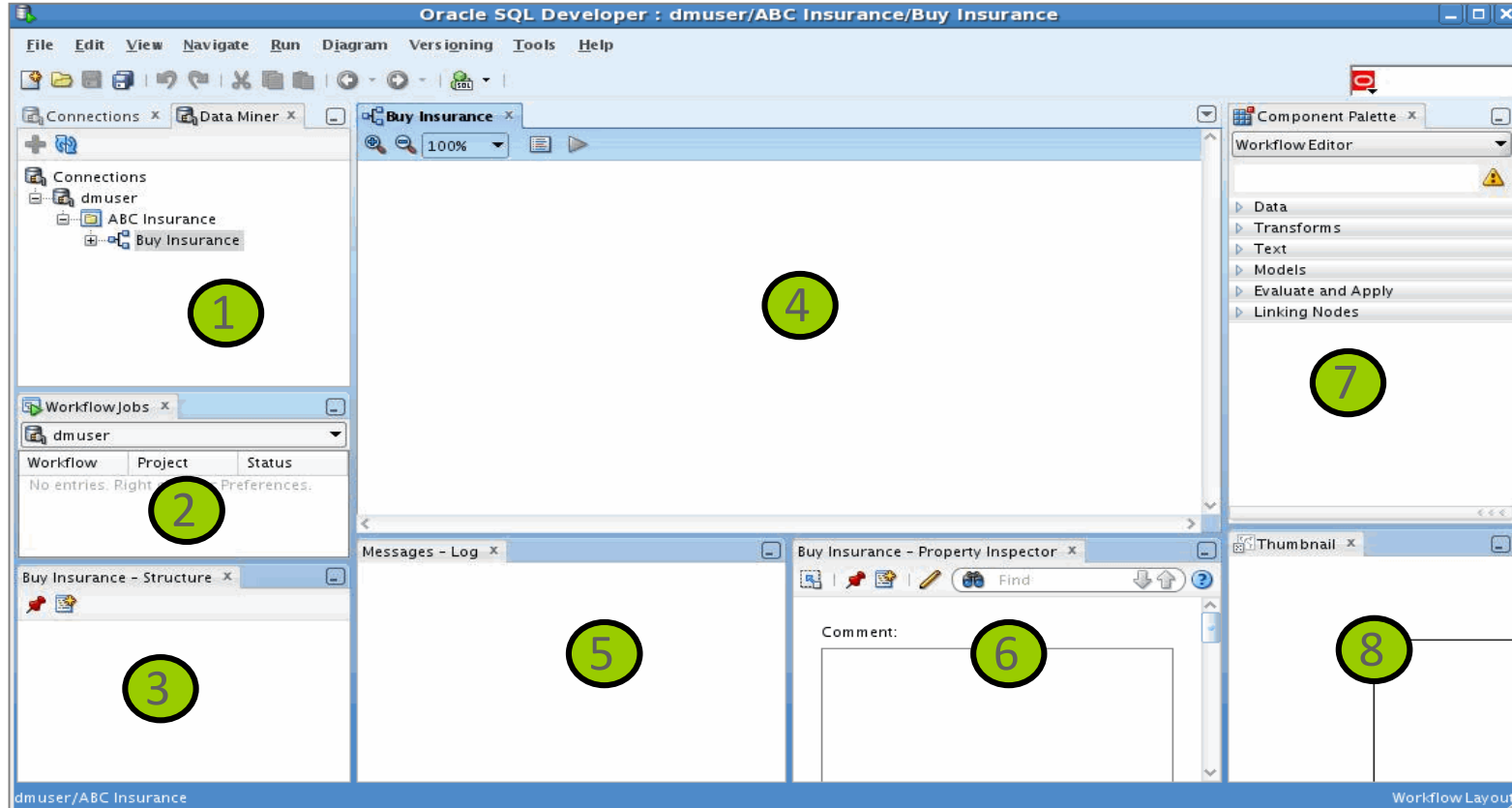
Running Task

Installing Schema

☒ Install Demo Data

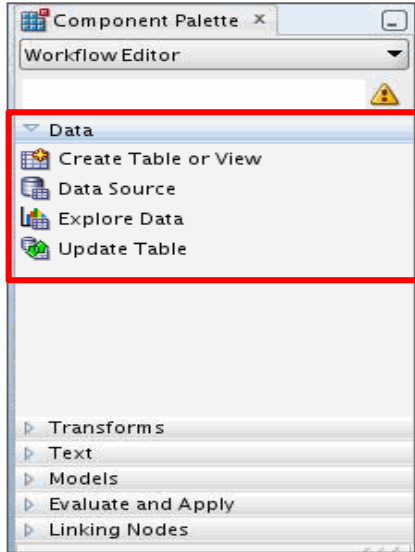
[Close](#)

Introducing the Data Miner Interface

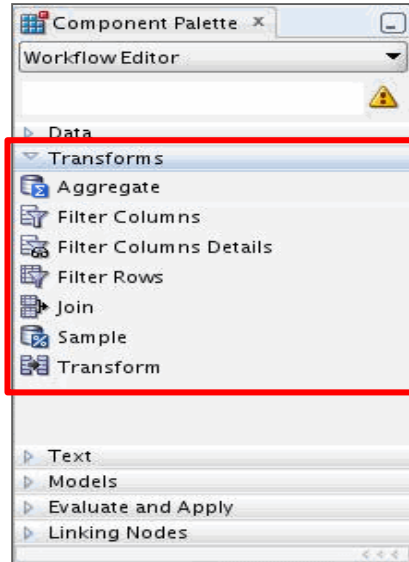


Examining Oracle Data Miner Nodes

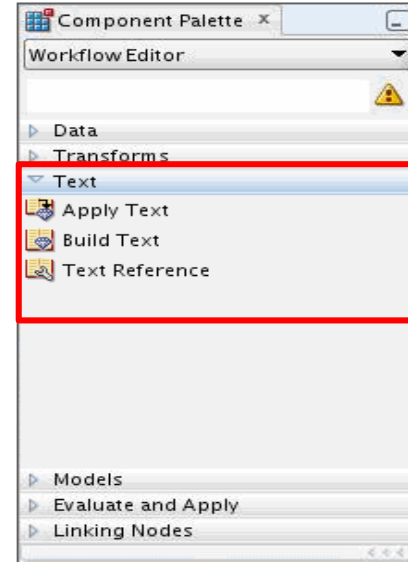
Data



Transforms

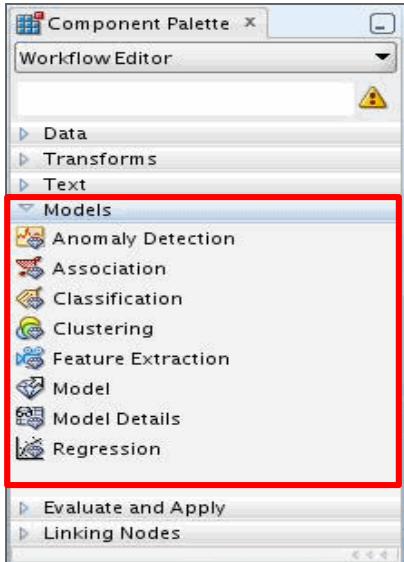


Text

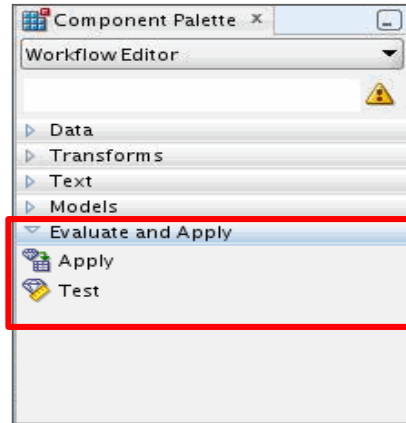


Examining Oracle Data Miner Nodes

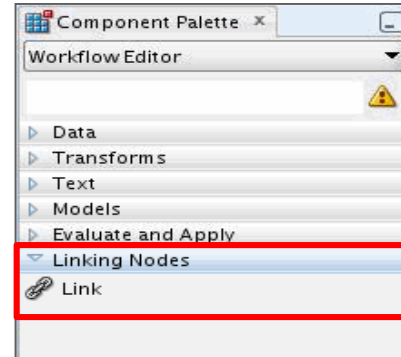
Models



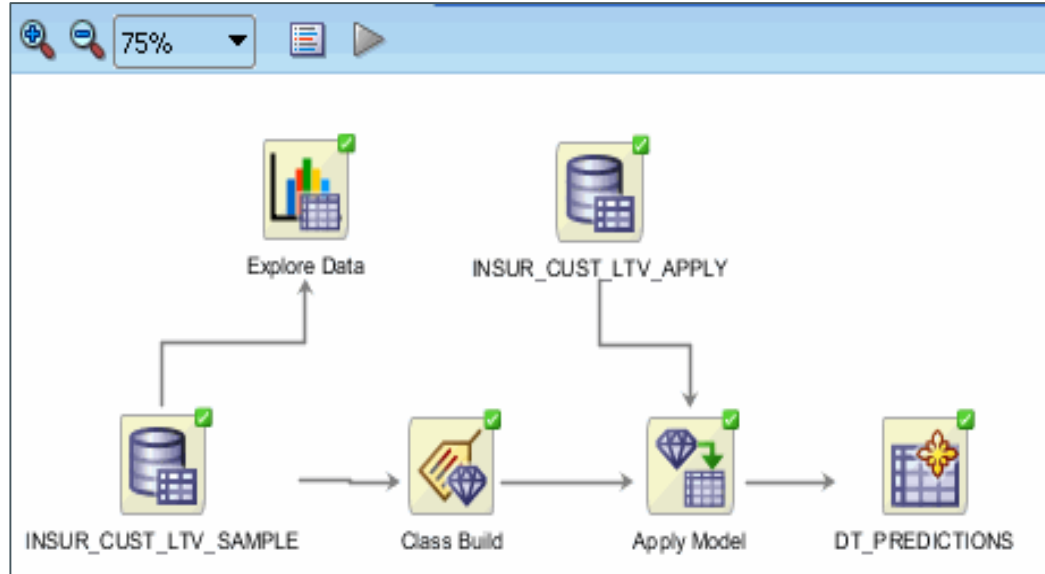
Evaluate and Apply



Linking



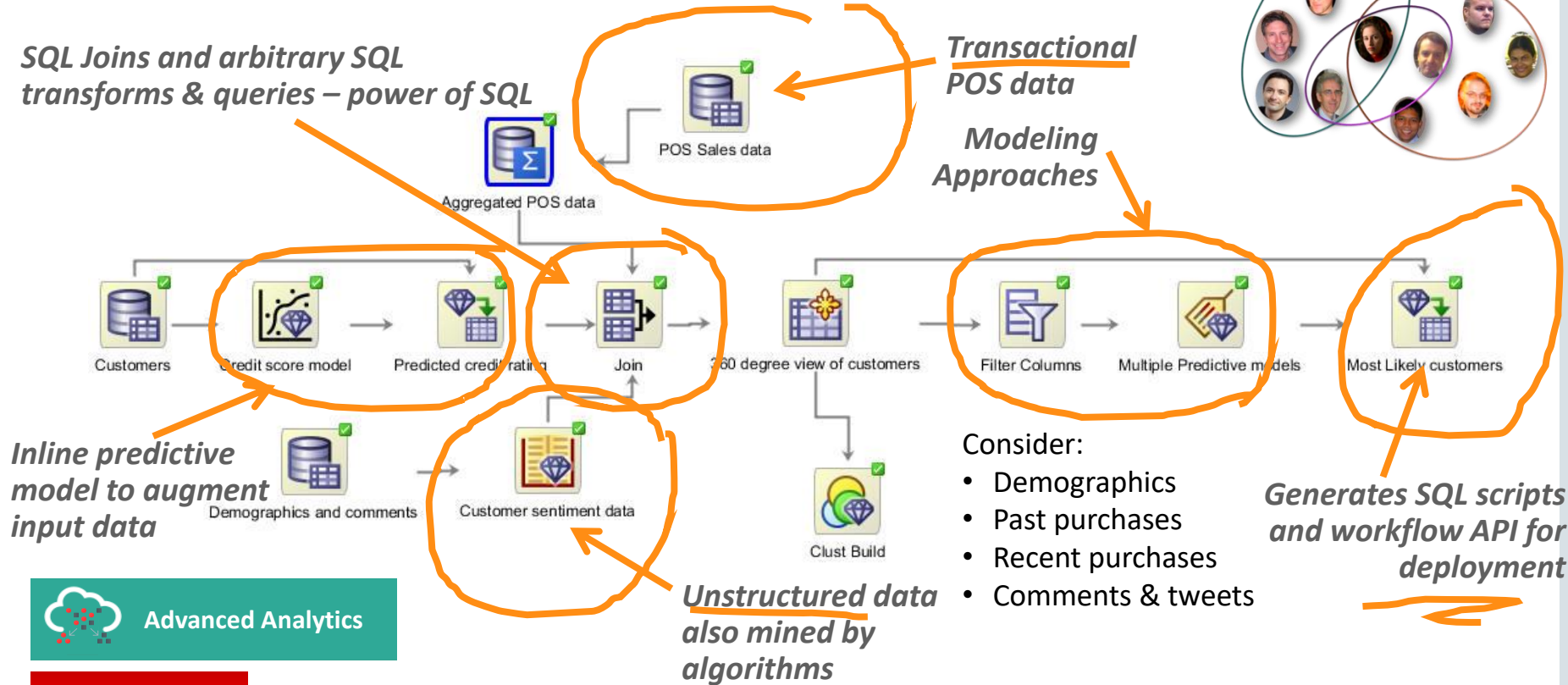
Previewing a Data Miner Workflow



Rapidly Build, Evaluate & Deploy Analytical Methodologies

Leveraging a Variety of Data Sources and Types

SQL Joins and arbitrary SQL transforms & queries – power of SQL



Advanced Analytics

ORACLE®

Previewing a 4.2 Feature

Workflow Scheduler

Oracle SQL Developer - jw415/Project/workflow

File Edit View Navigate Run Diagram Team Tools Window Help Automation

workflow - Structure Thumbnail Start Page workflow

Class Build INSUR_CUST_LTV_SAMPLE Links

Connections Data Miner

Connections Connections den00jff-m jw415 wfriday Oracle NetSQL Connections Cloud Connections

Workflow Jobs den00jff-m Workflow Project Status No entries. Right click for Preferences.

Class Build - Properties

Model Settings

Partition	Name	Output	Build	Test	Tune	Algorithm	Comment
Build	CLAS_SYM_1_1	→	1/15/16 7:15 AM	1/15/16 7:15 AM	Automatic	Support Vector Machine	
	CLAS_DT_1_1	→	1/15/16 7:15 AM	1/15/16 7:15 AM	Automatic	Decision Tree	
	CLAS_NB_1_1	→	1/15/16 7:15 AM	1/15/16 7:15 AM	Automatic	Naive Bayes	

Reports

All Reports

- Analytic View Reports
- Data Dictionary Reports
- Data Modeler Reports
- OLAP Reports
- TimesTen Reports
- User Defined Reports

Logging Page - Log

Level	Sequence	Elapsed	Source	Message
ERROR	2376	1155	o.dmt.dataminer.numme...	RunItem: unable to read repository config table.
ERROR	2375	1373	o.dmt.dataminer.numme...	RunItem: unable to read repository config table.
ERROR	2374	1513	o.dmt.dataminer.numme...	RunItem: unable to read repository config table.
ERROR	2373	1622	o.dmt.dataminer.numme...	RunItem: unable to read repository config table.
ERROR	2372	1092	o.dmt.dataminer.numme...	RunItem: unable to read repository config table.
ERROR	2371	1170	o.dmt.dataminer.numme...	RunItem: unable to read repository config table.

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Transforms

- Text
- Models
- Anomaly Detection
- Association
- Classification
- Clustering
- Explicit Feature Extraction
- Feature Extraction
- Model
- Model Details
- Regression
- Predictive Queries
- Model Operations
- Linking Nodes

den00jff-m

EN 2:56 PM 1/15/2016

Oracle SQL Developer: jw1415/Project1/workflow

File Edit View Navigate Run Diagram Team Tools Window Help Automation

workflow - Structure Thumbnail

Class Build

INSUR_CUST_LTV_SAMPLE

Links

Connections Data Miner

Connections

den00jf-im

jw1415

wfriday

Oracle NoSQL Connections

Cloud Connections

workflow - Properties

Comment:

Workflow Jobs

den00jf-im

Workflow Project Status

No entries. Right click for Preferences.

Reports

All Reports

Analytic View Reports

Data Dictionary Reports

Data Modeler Reports

OLAP Reports

TimesTen Reports

User Defined Reports

Logging Page - Log

Level	Sequence	Elapsed	Source	Message
✖	2376	1155	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2375	1373	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2374	1513	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2373	1622	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2372	1092	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2371	1170	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.

Messages Statements Logging Page

Components

Workflow Editor

Data

Create Table or View

Data Source

Explore Data

Graph

SQL Query

Update Table

Transforms

Text

Models

Anomaly Detection

Association

Classification

Clustering

Explicit Feature Extraction

Feature Extraction

Model

Model Details

Regression

Predictive Queries

Model Operations

Linking Nodes

INSUR_CUST_LTV_SAMPLE

Class Build

Schedule Workflow

Start Date: Jan 15, 2016 15:58

Repeat: Every Day

End Repeat: Never

☐ Use Existing Schedule

<Select a Schedule>

Help Advanced... OK Cancel

January 2016

Sun	Mon	Tue	Wed	Thu	Fri	Sat
27	28	29	30	31	1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31	1	2	3	4	5	6



workflow - Structure

- Class Build
- INSUR_CUST_LTV_SAMPLE
- Links

Connections - Data Miner

Connections

- den00jlf-im
- jw1415
- wfriday
- Oracle NoSQL Connections
- Cloud Connections

Workflow Jobs

den00jlf-im

Workflow	Project	Status
No entries. Right click for Preferences.		

Start Page

100%

Performance Options On (Selected)

INSUR_CUST_LTV_SAMPLE → Class Build

Schedule Workflow

Start Date: Jan 15, 2016 15:58

Repeat: Every Day

End Repeat: None

☐ Use Existing

<Select a Schedule>

Help Advanced... OK Cancel

workflow - Properties

Find

Comment:

Reports

- All Reports
- Analytic View Reports
- Data Dictionary Reports
- Data Modeler Reports
- OLAP Reports
- TimesTen Reports
- User Defined Reports

Logging Page - Log

Level	Sequence	Elapsed	Source	Message
ERROR	2376	1155	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
ERROR	2375	1373	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
ERROR	2374	1513	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
ERROR	2373	1622	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
ERROR	2372	1092	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
ERROR	2371	1170	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.

Messages Statements Logging Page

Components

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Transforms

Text

Models

- Anomaly Detection
- Association
- Classification
- Clustering
- Explicit Feature Extraction
- Feature Extraction
- Model
- Model Details
- Regression

Predictive Queries

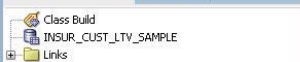
Model Operations

Linking Nodes

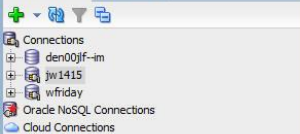




workflow - Structure Thumbnail



Connections Data Miner



Workflow Jobs

Workflow	Project	Status
No entries. Right click for Preferences.		

den00jlf-im

Start Page workflow

100%



Schedule Workflow

Start Date: Jan 15, 2016 15:58

Repeat: Every Week

End Repeat: Never

☐ Use Existing Schedule

<Select a Schedule>

Help Advanced... OK Cancel

workflow - Properties

Find

Comment:

Reports

- All Reports
- Analytic View Reports
- Data Dictionary Reports
- Data Modeler Reports
- OLAP Reports
- TimesTen Reports
- User Defined Reports

Logging Page - Log

Level	Sequence	Elapsed	Source	Message
✖	2376	1155	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2375	1373	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2374	1513	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2373	1622	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2372	1092	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2371	1170	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.

Messages Statements Logging Page

Components

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Advanced Settings

Notification Settings Nodes

Time Zone: America/New_York

Job Class: DEFAULT_JOB_CLASS

Job Priority: High Medium Low

☒ Max Failures: 3

☒ Max Run Duration: Days 1 Hours 0 Mins 0

☒ Schedule Limit: Days 0 Hours 2 Mins 0

Help OK Cancel

- association
- Classification
- Explicit Feature Extraction
- Feature Extraction
- Model
- Model Details
- Regression

- Predictive Queries
- Model Operations
- Linking Nodes

workflow - Structure Thumbnail

Start Page workflow

INSUR_CUST_LTV_SAMPLE

Class Build

INSUR_CUST_LTV_SAMPLE

Class Build

Schedule Workflow

Start Date: Jan 15, 2016 15:58

Repeat: Custom...

End Repeat: Never

☐ Use Existing Schedule

<Select a Schedule>

Help Advanced... OK Cancel

Repeat

Frequency: Weekly

Every 1 week(s) on:

☐ Sun ☒ Mon ☐ Tue ☐ Wed ☐ Thur ☐ Fri ☐ Sat

Help OK Cancel

Components

Workflow Editor

Data

Create Table or View

Data Source

Explore Data

Graph

SQL Query

Update Table

Connections Data Miner

Connections

den00jf-im

jw1415

wfriday

Oracle NoSQL Connections

Cloud Connections

Workflow Jobs

den00jf-im

Workflow Project Status

No entries. Right click for Preferences.

Reports

All Reports

Analytic View Reports

Data Dictionary Reports

Data Modeler Reports

OLAP Reports

TimesTen Reports

User Defined Reports

Logging Page - Log

Level	Sequence	Elapsed	Source	Message
✖	2376	1155	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2375	1373	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2374	1513	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2373	1622	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2372	1092	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.
✖	2371	1170	o.dmt.dataminer.runne...	RunItem: unable to read repository config table.

Messages Statements Logging Page

Models

Anomaly Detection

Association

Classification

Clustering

Explicit Feature Extraction

Feature Extraction

Model

Model Details

Regression

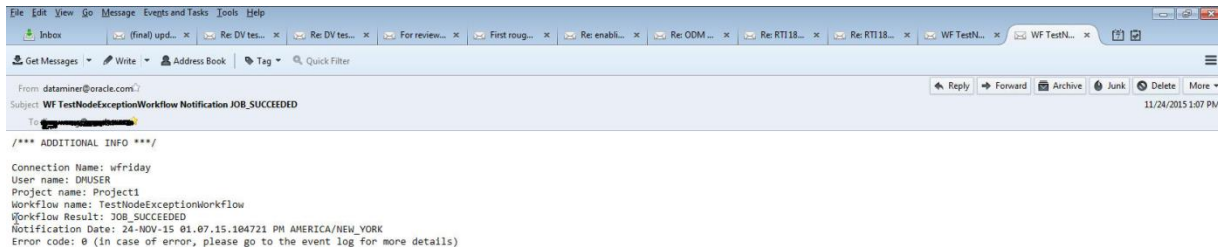
Predictive Queries

Model Operations

Linking Nodes

Previewing a 4.2 Feature

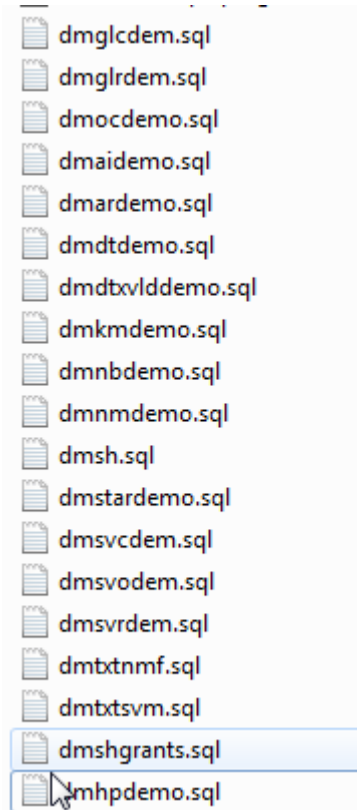
Workflow Scheduler—Email Notifications



The Data Mining Sample Programs

- The [Data Mining Sample Programs 12c Documentation](#)

- You can learn a great deal about the Oracle Data Mining API from the data mining sample programs. The programs illustrate typical approaches to data preparation, algorithm selection, algorithm tuning, testing, and scoring.
- The programs are easy to use. They include extensive inline comments to help you understand the code. They delete all temporary objects on exit; you can run the programs repeatedly without setup or cleanup.
- The data mining sample programs are installed with Oracle Database Examples in the demo directory under Oracle Home. The demo directory contains sample programs that illustrate many features of Oracle Database. You can locate the data mining files by doing a directory listing of dm*.sql.



The Data Mining Sample Programs

Attribute Importance Sample Code

```
Worksheet Query Builder
5.09600019 seconds

-- BUILD THE MODEL

-----

-- Cleanup old output table for repeat runs
BEGIN EXECUTE IMMEDIATE 'DROP TABLE ai_explain_output';
EXCEPTION WHEN OTHERS THEN NULL; END;
/

-----

-- Run the EXPLAIN routine to get attribute importance results
BEGIN
  DBMS_PREDICTIVE_ANALYTICS.EXPLAIN(
    data_table_name => 'mining_data_build_v',
    explain_column_name => 'affinity_card',
    result_table_name => 'ai_explain_output');
END;
/

-----

-- DISPLAY RESULTS
```

Script Output x

Task completed in 5.096 seconds

ATTRIBUTE_NAME	EXPLANATORY_VALUE	RANK
HOUSEHOLD_SIZE	.195	1
CUST_MARITAL_STATUS	.194	2
YRS_RESIDENCE	.115	3
EDUCATION	.106	4
AGE	.104	5
OCCUPATION	.092	6
Y_BOX_GAMES	.077	7
HOME_THEATER_PACKAGE	.069	8
CUST_GENDER	.043	9
BOOKKEEPING_APPLICATION	.024	10
BULK_PACK_DISKETTES	.000	11
COUNTRY_NAME	.000	11
CUST_ID	.000	11
CUST_INCOME_LEVEL	.000	11
FLAT_PANEL_MONITOR	.000	11
OS_DOC_SET_KANJI	.000	11
PRINTER_SUPPLIES	.000	11

17 rows selected

Learn Predictive Analytics in 2 Hours!

Oracle Advanced Analytics Hands on Lab

- **READY, SET, GO!!!!**

- Recommend doing 3-5 Tutorials

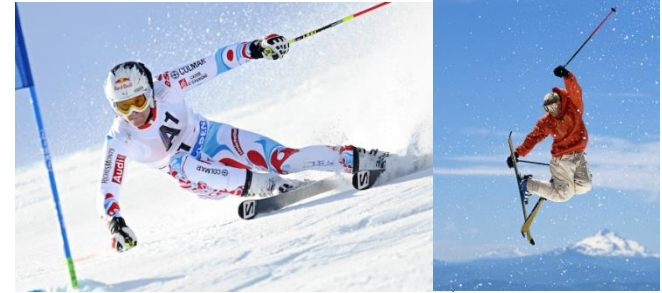
1. Using Oracle Data Miner 4.1

2. Star Schema Mining Using Oracle Data Miner 4.1

3. Text Mining with an EM Clustering Model Using Data Miner 4.1

4. **Anomaly Detection (CLAIMS)** *See Instructor for assistance*

5. **Market Basket Analysis (SH.SALES)** *See Instructor for assistance*



Novice/Introductory/Overviews

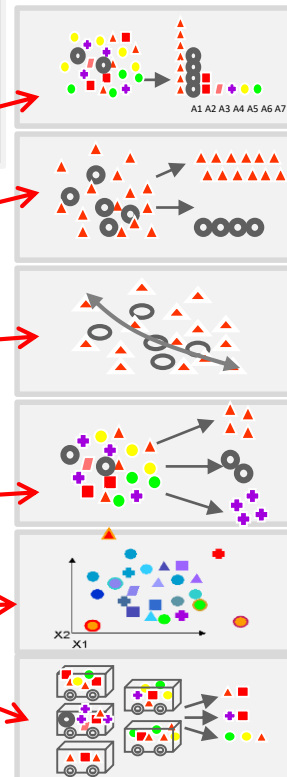
Quick Overview of Concepts, Process and Use Cases



What is Machine Learning, Data Mining & Predictive Analytics?

Automatically sifting through **large amounts** of data to create models that **find previously hidden patterns**, **discover valuable new insights** and **make predictions**

- Identify most important factor (*Attribute Importance*)
- Predict customer behavior (*Classification*)
- Predict or estimate a value (*Regression*)
- Find profiles of targeted people or items (*Decision Trees*)
- Segment a population (*Clustering*)
- Find fraudulent or “rare events” (*Anomaly Detection*)
- Determine co-occurring items in a “baskets” (*Associations*)



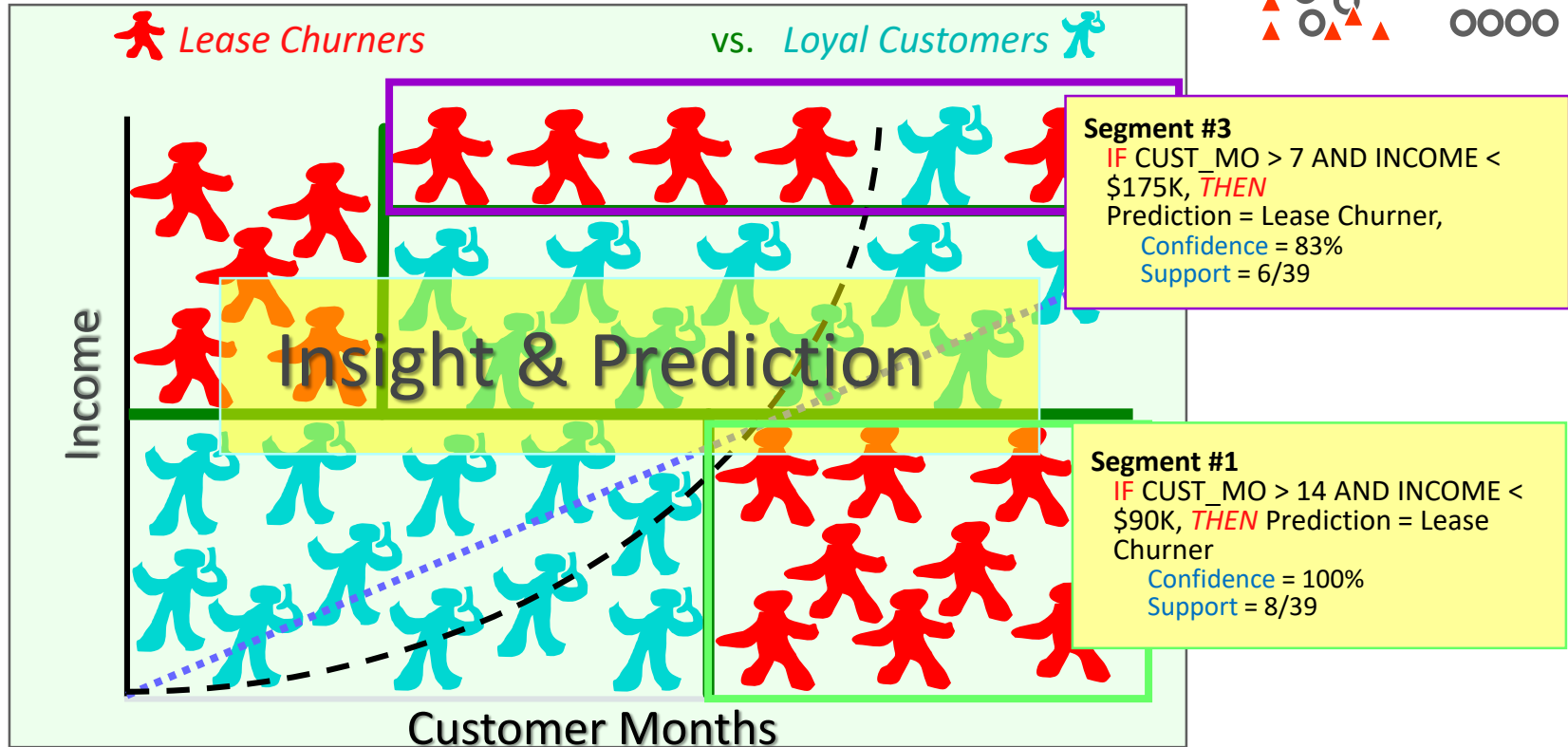
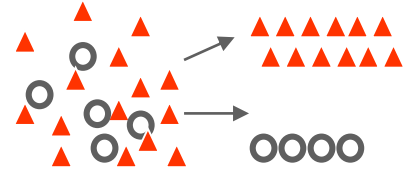
Predictive Analytics & Data Mining

Typical Use Cases

- Targeting the right customer with the right offer
- How is a customer likely to respond to an offer?
- Finding the most profitable growth opportunities
- Finding and preventing customer churn
- Maximizing cross-business impact
- Security and suspicious activity detection
- Understanding sentiments in customer conversations
- Reducing medical errors & improving quality of health
- Understanding influencers in social networks



Data Mining Provides Better Information, Valuable Insights and Predictions



Oracle Advanced Analytics DB Option

In-Database Machine Learning Algorithms*—SQL &  & GUI Access



Advanced Analytics



Classification



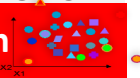
- Decision Tree
- Logistic Regression (GLM)
- Naïve Bayes
- Support Vector Machine (SVM)
- Random Forest

Regression



- Multiple Regression (GLM)
- Support Vector Machine (SVM)
- Stepwise Linear Regression
- Linear Model
- Generalized Linear Model
- Multi-Layer Neural Networks

Anomaly Detection



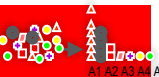
- 1-Class Support Vector Machine

Clustering



- Hierarchical k-Means
- Orthogonal Partitioning Clustering
- Expectation-Maximization

Attribute Importance



- Minimum Description Length
- Unsupervised pair-wise KL div.

Market Basket Analysis



- Apriori – Association Rules

Text Mining



- All OAA/ODM SQL ML support
- Explicit Semantic Analysis

Predictive Queries



- Clustering
- Regression
- Anomaly Detection
- Feature Extraction

Feature Extraction & Creation

- Nonnegative Matrix Factorization
- Principal Component Analysis
- Singular Value Decomposition

Time Series

- Single & Double Exp. Smoothing

Open Source R Algorithms

- Ability to run any R package (9,000+) via Embedded R mode



+ Ability to Mine Unstructured, Structured & Transactional data
+ Partitioned Models

ORACLE

Oracle University's Learn Predictive Analytics Using Oracle Data Mining Course Agenda

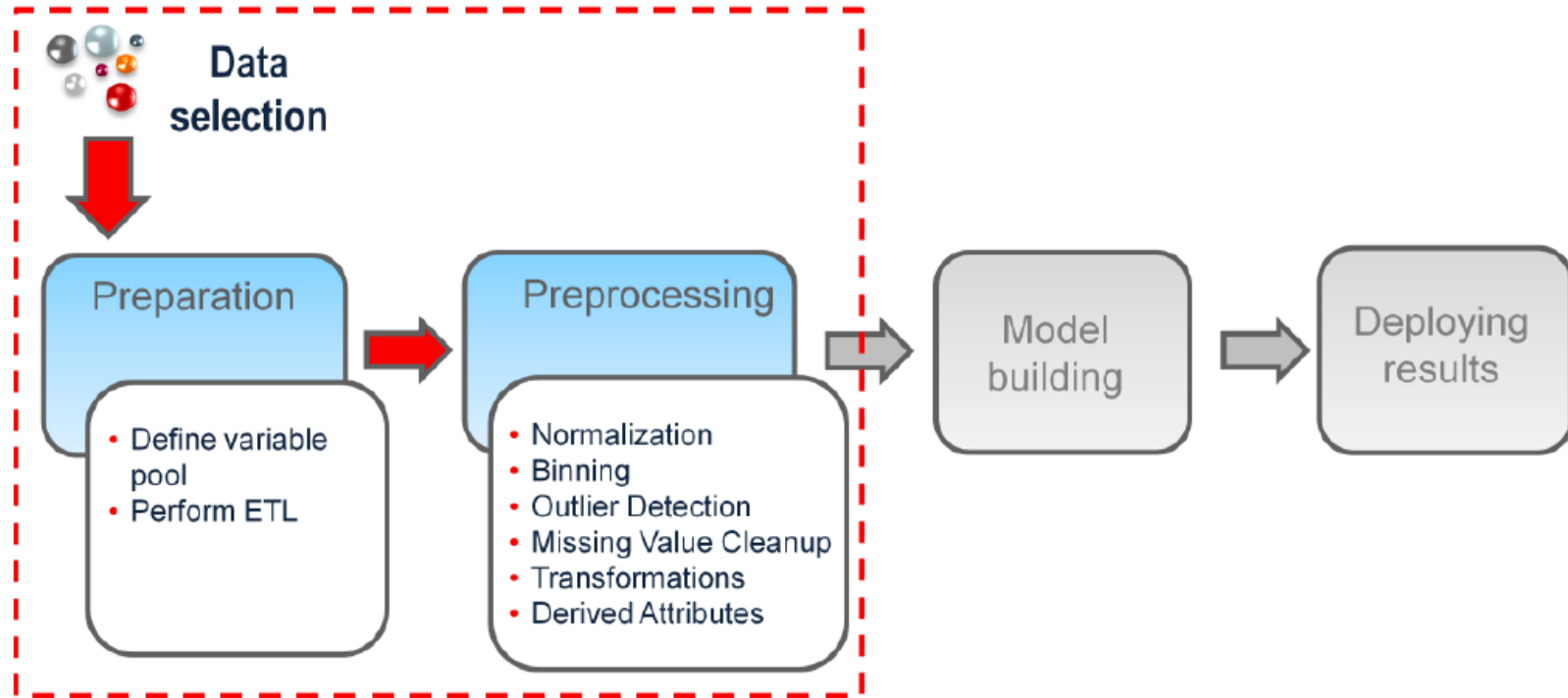
Day 1

1. Introduction
2. Predictive Analytics and Data Mining Concepts
3. The Data Mining Process
4. Introducing Oracle Data Miner 4.1
5. Using Classification Models
6. Using Regression Models

Day 2

7. Using Clustering Models
8. Performing Market Basket Analysis
9. Performing Anomaly Detection
10. Performing Star Schema Mining
11. Using Predictive Queries
12. Deploying Data Mining Results

The Data Mining Process



Data Mining Attributes

Data mining attributes can be:

- Categorical
- Numerical
- Text
- Target
- Active or Inactive

OCCUPATION
Prof.
Sales
Machine
Crafts

Categorical

AGE
40
47
32
63
40

Numerical

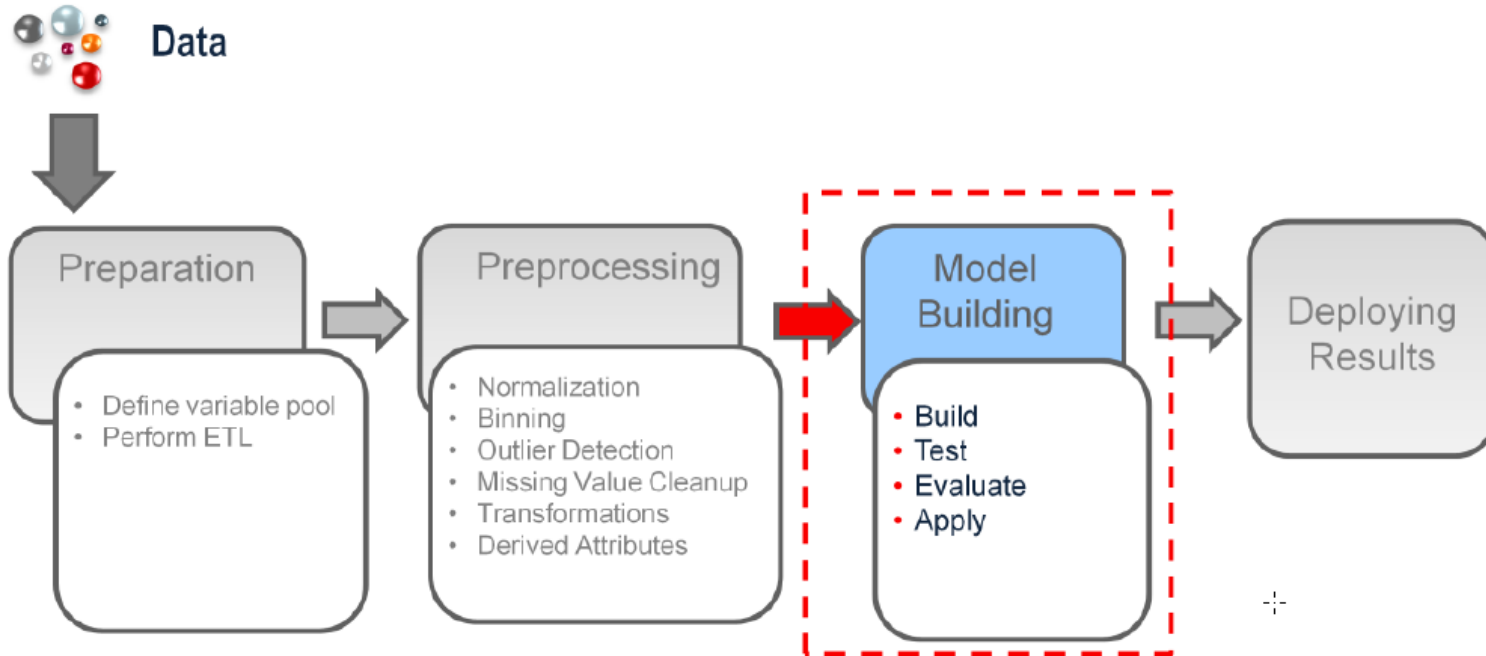
CLASS
1
0
0
0
0

Target

EDUCATION
Bach.
5th-6th
HS-grad
Bach.
HS-grad

Active

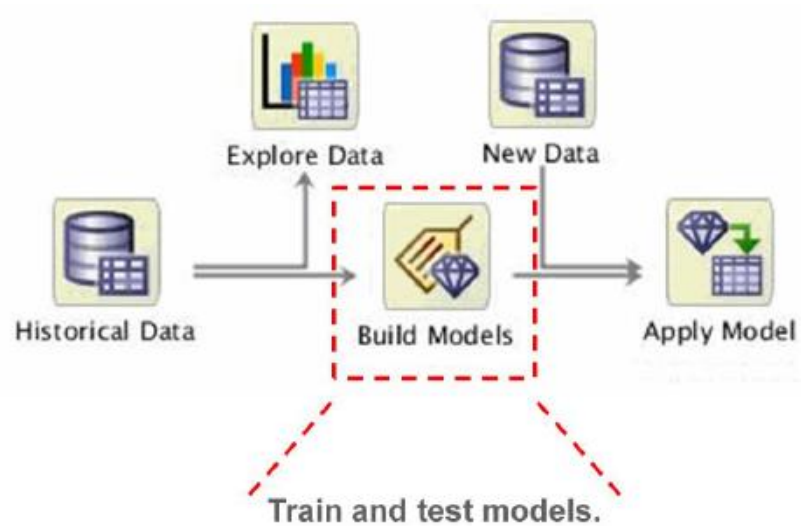
Building and Evaluating Models



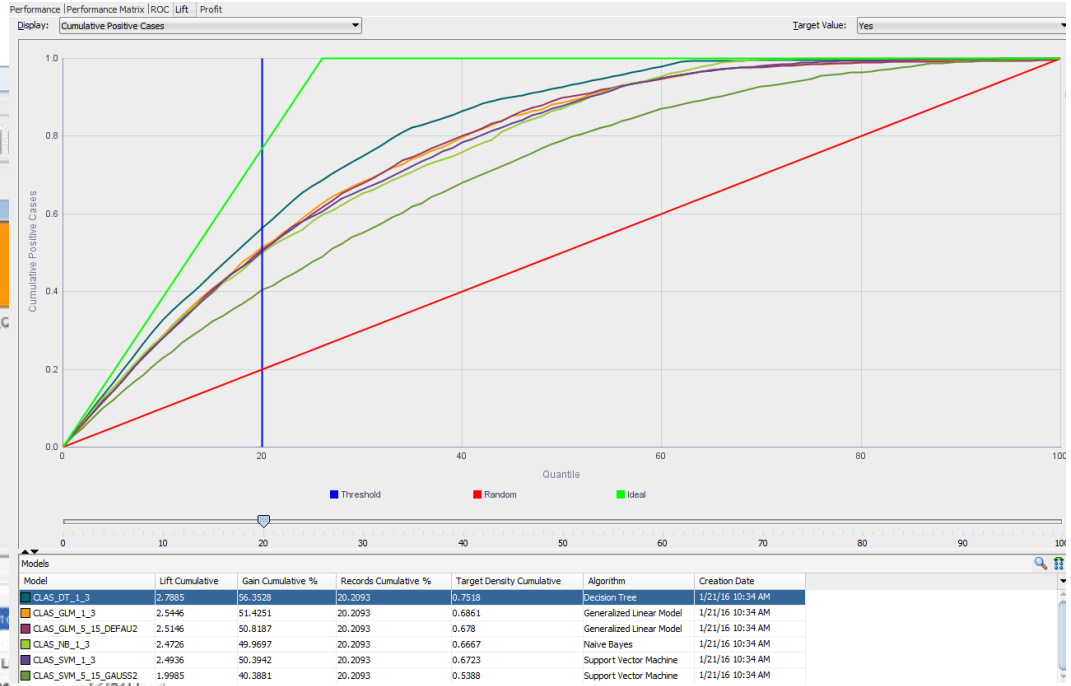
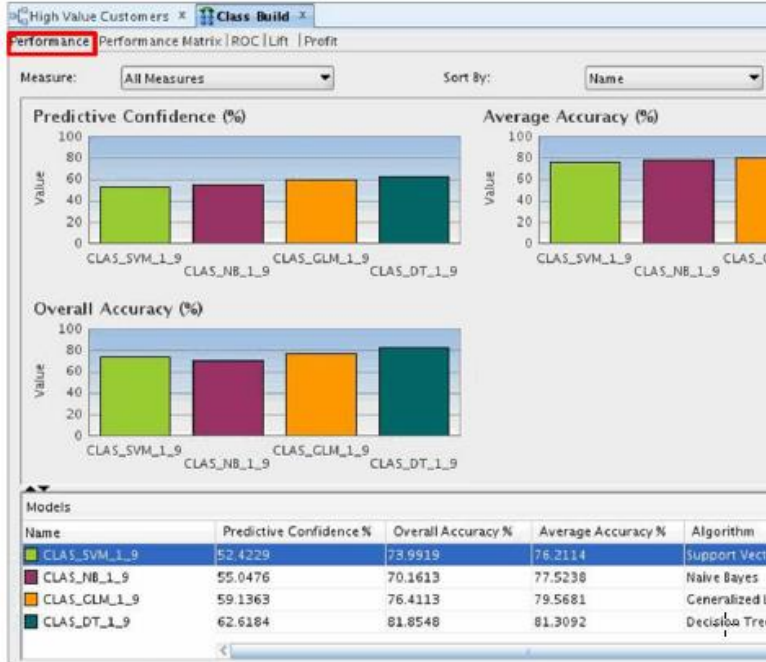
Model Building Tasks



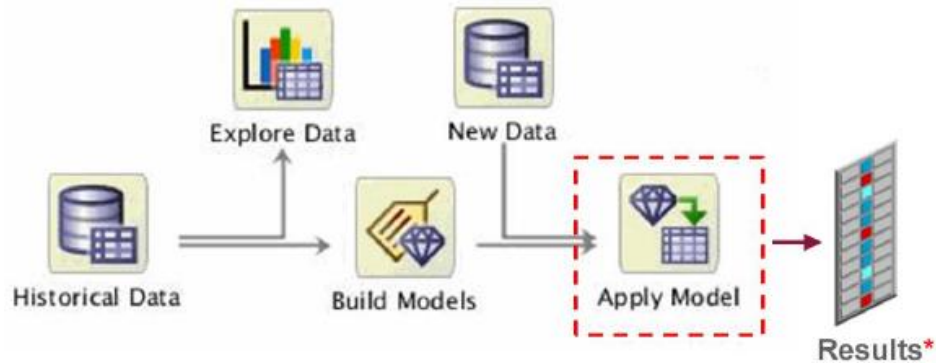
Model Train & Test: Supervised Learning



Model Evaluation: Supervised Learning



Applying the Selected Model(s)



(* Output settings may be specified by the user.)

A woman with long brown hair and black-rimmed glasses is seated at a wooden desk in a bright, modern office. She is wearing a brown leather jacket over a blue and black patterned scarf. She is holding a black smartphone to her ear with her left hand and looking down at a large open book or portfolio on the desk with her right hand. In the background, another person is seated at a desk, and large windows let in natural light.

Oracle Advanced Analytics

Brief Demos

ORACLE®

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

FileEditViewRunDiagramTeamToolsWindowHelp

ConnectionsData Miner

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY_INSURANCE_WORKFLOW

CARS_DATA_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

OOO14_CUSTOMER_ANALYTICS_360_I

Predictive Oueries

Thumbnail

BUY_INSURANCE_WORKFLOW - ...

Reports

Clustering Segmentation

Explore/Profile Data

Prediction Queries by Region

Prediction Query

Filter Cols_Attrib Importance

Predictive Models for Student

Class Build

Predictive models ind TEXT

Filter Columns Details

Scatter Box plots etc.

Graph

Anomaly Detection Query

CUST_INSUR_LTV

CUST_INSUR_LTV_APPLY

MINING_DATA_TEXT_BUILD_V

Start Page

BUY_INSURANCE_WORKFLOW

100%

Parallel Query Off

CUST_INSUR_LTV1

CUST_INSUR_LTV1 - Properties

Find

Data

Cache

Details

Source Table: DMUSER.CUST_INSUR_LTV

Name

Alias

Data Type

AGE

NUMBER

BANK_FUNDS

NUMBER

BUY_INSURANCE

VARCHAR2

CAR_OWNERSHIP

NUMBER

CHECKING_AMOUNT

NUMBER

Components

Workflow Editor

Data

Create Table or View

Data Source

Explore Data

Graph

SQL Query

Update Table

Transforms

Aggregate

Filter Columns

Filter Columns Details

Filter Rows

Join

JSON Query

Sample

Transform

Text

Models

Anomaly Detection

Association

Classification

Clustering

Feature Extraction

Model

Model Details

Regression

Predictive Queries

Evaluate and Apply

Linking Nodes

File Edit View Navigate Run Diagram Team Tools Window Help

Connections Data Miner Start Page BUY_INSURANCE_WORKFLOW

Connections

- dmuser
 - ACME Mfg Paint Project
 - BERGERS R US
 - Fun with Gov
 - Predictive Analytics WF
 - Chicago Crime
 - Customers R Us Project
 - A+ Students OAA analytics
 - Big Data Analytics w JSON
 - Big Data Analytics
 - BUY_INSURANCE_WORKFLOW
 - CARS_DATA_MINING
 - Churners01 work flow
 - Claims Fraud Clustering + SVM2
 - Customer Analytics ind RFM
 - Employees_attrition
 - Insurance Customer Analytics
 - Manufacturing Painting Ops
 - Market Basket Analysis
 - OOOW14_CUSTOMER_ANALYTICS_360_I
 - Predictive Queries

Thumbnail

BUY_INSURANCE_WORKFLOW - ... Reports

Clustering Segmentation

Explore/Profile Data

Prediction Queries by Region

Prediction Query

Filter Cols_Attrb Importance

Predictive Models for Student

Class Build

Predictive models ind TEXT

Filter Columns Details

Scatter Box plots etc.

Graph

Anomaly Detection Query

CUST_INSUR_LTV

CUST_INSUR_LTV_APPLY

MINING_DATA_TEXT_BUILD_V

Connect

Run

Force Run

Edit...

View Data

Generate Apply Chain

Show Event Log

Validate Parents

Deploy

Save SQL

Cut Ctrl-X

Copy Ctrl-C

Paste Ctrl-V

Extended Paste... Ctrl+Shift-V

Select All Ctrl-A

Parallel Query ...

Copy Image to Clipboard

Save Image As...

Go to Properties

CUST_INSUR_LTV1 - Properties

Find

Data

Cache

Details

Source Table: DMUSER.CUST_INSUR_LTV

Name	Alias	Data Type
AGE		NUMBER
BANK_FUNDS		NUMBER
BUY_INSURANCE		VARCHAR2
CAR_OWNERSHIP		NUMBER
CHECKING_BALANCE		NUMBER

Components

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Transforms

- Aggregate
- Filter Columns
- Filter Columns Details
- Filter Rows
- Join
- JSON Query

Text

Models

- Anomaly Detection
- Association
- Classification
- Clustering
- Feature Extraction
- Model
- Model Details
- Regression

Predictive Queries

Evaluate and Apply

Linking Nodes

dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

CUST_ID														
N_TRANS_ATM														
CHECKING_AMOUNT														
MARITAL_STATUS														
SEX														
N_TRANS_KIOSK														
STATE														
HOUSE_OWNERSHIP														
MONTHLY_CHECKS_WRITTEN														
LTV														
SALARY														
BANK_FUNDS														
BUY_INSURANCE														
CREDIT_BALANCE														
N_MORTGAI														
1	CU2404	6	25	WIDOWED	F	9	DC	1	2	25,370...	63,883	25,000	No	549
2	CU2405	0	25	MARRIED	M	2	MI	1	0	29,594.5	59,978	0	Yes	0
3	CU2406	5	25	WIDOWED	M	2	CA	1	4	26,473.5	67,894	17,200	Yes	13,859
4	CU2407	2	3,226	DIVORCED	F	2	MI	1	2	25,976...	67,107	7,600	No	0
5	CU2408	4	130	DIVORCED	F	2	NY	1	11	28,446.5	70,986	1,500	No	23,525
6	CU2409	4	6,008	DIVORCED	F	2	CA	0	2	13,921...	58,887	10,900	No	0
7	CU2411	4	10,943	MARRIED	M	2	NY	1	2	23,833...	69,335	0	No	4,734
8	CU2412	5	25	MARRIED	F	2	FL	1	12	22,485.5	58,342	2,600	Yes	0
9	CU2413	4	293	SINGLE	M	1	NY	0	2	18,865...	67,061	1,614	Yes	2,249
10	CU2414	2	25	DIVORCED	M	2	CA	1	2	29,672...	67,489	0	No	0
11	CU2416	2	25	SINGLE	M	2	NY	0	4	20,190...	63,963	0	No	0
12	CU2417	2	2,420	SINGLE	M	2	NY	0	4	19,941	64,964	1,156	No	963
13	CU2418	2	1,089	DIVORCED	F	2	MI	1	3	27,892...	58,771	6,850	No	0
14	CU2420	5	25	DIVORCED	M	2	NY	1	3	19,536.5	60,946	5,400	No	44,020
15	CU2421	5	882	DIVORCED	F	2	CA	1	4	29,123.5	67,694	3,950	No	0
16	CU2422	3	157	MARRIED	F	2	MI	1	4	29,145...	61,781	1,850	Yes	0
17	CU2423	4	9,894	WIDOWED	F	8	NY	1	12	25,868...	63,875	3,300	No	6,707
18	CU3100	6	227	WIDOWED	F	2	MI	1	9	17,539...	65,359	4,280	No	0
19	CU3101	4	40	MARRIED	F	2	MI	1	11	23,178...	61,513	800	Yes	0
20	CU3102	4	21,094	SINGLE	M	1	NV	0	4	21,644...	65,379	8,129	No	0
21	CU3104	5	25	MARRIED	F	2	NY	1	3	16,525.5	66,902	17,001	No	92,408
22	CU3105	7	638	OTHER	F	2	NY	1	3	24,184.5	63,538	20,200	Yes	0
23	CU3106	4	372	DIVORCED	M	2	MI	1	12	26,580.5	69,522	0	No	0
24	CU3107	0	25	SINGLE	M	2	CA	0	0	19,374	64,296	0	No	0
25	CU3108	3	25	MARRIED	F	2	NY	1	4	14,843...	60,575	700	Yes	0
26	CU3109	4	25	MARRIED	M	2	MI	1	1	30,839.5	74,158	0	No	0
27	CU3110	4	25	MARRIED	F	2	CA	1	2	21,882...	59,929	2,300	No	36,102
28	CU3111	4	25	DIVORCED	M	2	MN	1	7	27,128...	55,713	3,650	No	0
29	CU3113	1	25	SINGLE	M	2	MS	0	1	15,834.5	64,138	0	No	0
30	CU3115	2	25	MARRIED	M	3	CA	1	3	27,130	66,120	1,600	No	0
31	CU3116	4	17,173	DIVORCED	F	2	CA	1	18	24,554	68,616	4,000	No	11,098
32	CU3117	4	13,225	SINGLE	M	1	MI	0	11	16,890	65,560	1,606	No	0
33	CU3118	4	7,543	SINGLE	M	3	CA	1	1	28,260...	71,843	700	No	0
34	CU3119	4	25	DIVORCED	F	3	MI	1	3	27,410	58,840	0	No	0
35	CU3120	2	25	SINGLE	M	3	CA	0	1	0	63,095	0	No	0
36	CU3121	4	543	DIVORCED	M	3	CA	1	4	28,178...	60,315	5,700	No	0
37	CU3123	1	25	SINGLE	M	3	NY	0	1	24,367.5	72,670	0	No	0
38	CU3124	4	12,988	SINGLE	M	3	MN	0	2	17,477.5	70,310	10,100	No	0
39	CU3125	6	25	DIVORCED	M	3	CA	1	16	22,995	68,380	5,950	No	66,417
40	CU3126	2	108	DIVORCED	F	3	MI	1	9	24,825	60,900	7,200	No	0

FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData Miner

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY_INSURANCE_WORKFLOW

CARS_DATA_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

OOW14_CUSTOMER_ANALYTICS_360_I

Thumbnail

BUY_INSURANCE_WORKFLOW

Reports

Class Build

Predictive models ind TEXT

Filter Columns Details

Scatter Box plots etc.

Graph

Anomaly Detection Query

CUST_INSUR_LTV

CUST_INSUR_LTV_APPLY

MINING_DATA_TEXT_BUILD_V

Model Details

GLM std coefficients

PREDICTIONS

OUTPUT_12_15

Likely Buyers

Customer Segments

CUST_INSUR_LTV1

Explore Data

Start Page

BUY_INSURANCE_WORKFLOW

CUST_INSUR_LTV1

100%

Parallel Query Off

CUST_INSUR_LTV1

Explore Data

Components

Workflow Editor

Data

Create Table or View

Data Source

Explore Data

Graph

SQL Query

Update Table

Transforms

Aggregate

Filter Columns

Filter Columns Details

Filter Rows

Join

JSON Query

Text

Models

Anomaly Detection

Association

Classification

Clustering

Feature Extraction

Model

Model Details

Regression

Predictive Queries

Evaluate and Apply

Linking Nodes

Explore Data - Properties

Find

Input

Statistics

Output

Histogram

Sample

Details

Group By:

<Select Group By>

☒ Auto Input Columns Selection

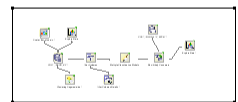
Name	Data Type
AGE	NUMBER
BANK_FUNDS	NUMBER
BUY_INSURANCE	VARCHAR2



Connections Data Miner

- Connections
- BIWA16_12c_DB_Cloud_Jonathan2
 - BIWA DMUSER - Learn Predictive Analytics in 2 Hours HO
 - CBERGER dbpm23 Cloud instance Connection
 - Charlie Database 12c laptop dmuser_dmuser
 - AAA Customer Analytics
 - 360 View Star Schema Analytics
 - BUY INSURANCE
 - Claims Anomaly Detection
 - Customer 360 Clustering
 - Customer analytics NEW
 - Fun Analytics WF
 - Market Basket Analysis
 - My First Workflow
 - New Buy Insurance WF
 - New Customer Analytics
 - New WF
 - Predictive Queries
 - RFM Input4 Customer Analytics
 - Structured Unstructured Data

Thumbnail



SSH Hosts

- SSH Hosts
- CBERGER dbpm23 Cloud instance
 - Express Cloud PDB May 25
 - HOL3
 - NEW_SSH_BIWA_Brian2
 - OracleCloudDatabase1
 - OracleDatabaseCloud2
 - SSH_BIWA_Jonathan2

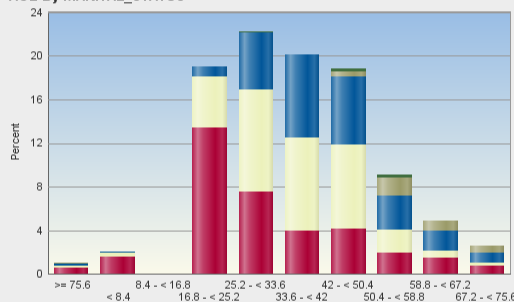
Start Page BUY INSURANCE Charlie Database 12c laptop dmuser_dmuser2.sql Profile Data

Statistics Data Columns SQL

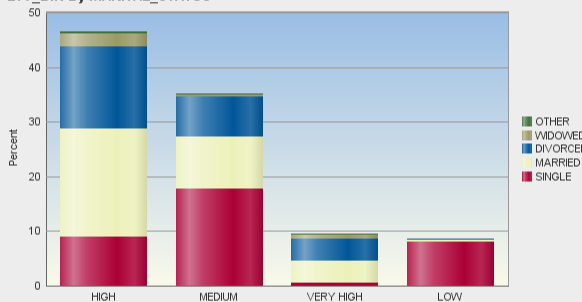
Statistics: 10 Columns from 2,005 Rows(Sampled)

Name	Histogram	Data Type	Percent NULLS	Distinct Values	Distinct Per...	Mode	Average	Median	Min Value	Max Value	Standard Devi...	Variance
AGE		NUMBER	0	69	3.4414		37.6823	36	0	84	14.4737	209.4873
BANK_FUNDS		NUMBER	0	425	21.197		2,585.3736	551	0	30,200	4,739.2687	22,460,667.7...
BUY_INSURANCE		VARCHAR2	0	2	0.0998	No						
CAR_OWNERSHIP		NUMBER	0	2	0.0998		0.9247	1	0	1	0.264	0.0697
CHECKING_AMOUNT		NUMBER	0	626	31.2219		1,032.7312	25	25	24,471	3,103.6812	9,632,836.6997

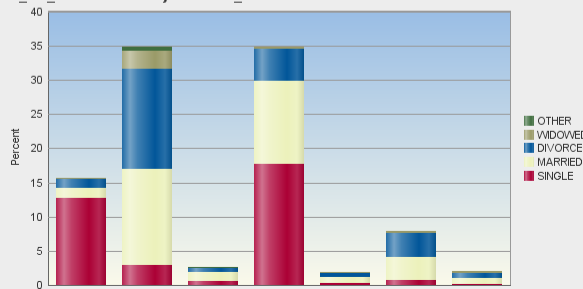
AGE By MARITAL_STATUS



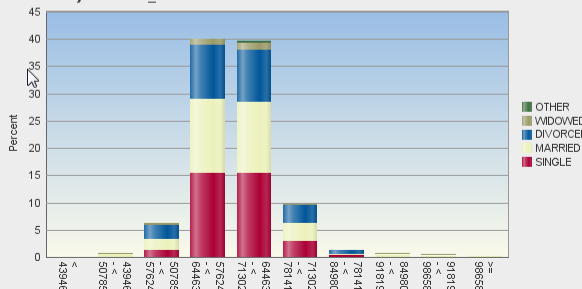
LTV_BIN By MARITAL_STATUS



N_OF_DEPENDENTS By MARITAL_STATUS



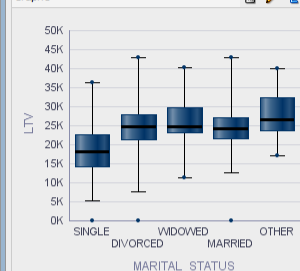
SALARY By MARITAL_STATUS



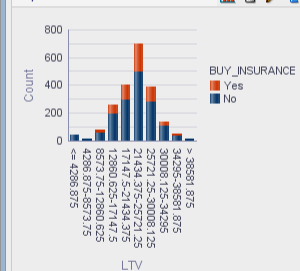
Scatter Box plots etc. Components

View: Sample Data Parallel Query Off...

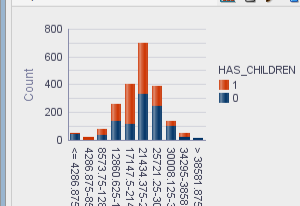
Graph1



Graph2



Graph3



Connections

- dmuser
 - ACME Mfg Paint Project
 - BERGERS R US
 - Fun with Gov
 - Predictive Analytics WF
 - Chicago Crime
 - Customers R Us Project
 - A+ Students OAA analytics
 - Big Data Analytics w JSON
 - Big Data Analytics
 - BUY_INSURANCE_WORKFLOW**
 - CARS_DATA_MINING
 - Churners01 work flow
 - Claims Fraud Clustering + SVM2
 - Customer Analytics ind RFM
 - Employees_attrition
 - Insurance Customer Analytics
 - Manufacturing Painting Ops
 - Market Basket Analysis

Thumbnail

BUY_INSURANCE_WORKFLOW - ...

- Predictive Models for Student
- Class Build
- Predictive models ind TEXT
- Filter Columns Details
- Scatter Box plots etc.
- Graph
- Graph 1
- Anomaly Detection Query
- CUST_INSUR_LTV
- CUST_INSUR_LTV_APPLY
- MINING_DATA_TEXT_BUILD_V
- CUST_INSUR_LTV1
- Model Details
- GLM std coefficients
- PREDICTIONS
- OUTPUT_12_15
- Likely Buyers
- Customer Segments
- Filter Columns

Start Page | BUY_INSURANCE_WORKFLOW | CUST_INSUR_LTV1 | Graph 1

100%

Parallel Query Off

Filter Columns - Properties

Filter: Find

Columns	Name	Type	Output	Hint
AGE	AGE	NUMBER	→	
BANK_FUNDS	BANK_FUNDS	NUMBER	→	
BUY_INSURANCE	BUY_INSURANCE	VARCHAR2	→	
CAR_OWNERSHIP	CAR_OWNERSHIP	NUMBER	→	
CHECKING_AMOUNT	CHECKING_AMOUNT	NUMBER	→	

Filter

- Connect
- Run
- Force Run**
 - Selected Node
 - Selected Node and Children
 - Selected Node and Parents
 - Child Nodes Only
- Edit...
- View Data
- View Attribute Importance
- Generate Apply Chain
- Show Event Log
- Validate Parents
- Deploy
- Save SQL
- Cut (Ctrl-X)
- Copy (Ctrl-C)
- Paste (Ctrl-V)
- Extended Paste... (Ctrl+Shift-V)
- Select All (Ctrl-A)
- Parallel Query ...
- Copy Image to Clipboard
- Save Image As...
- Go to Properties
- Navigate

Components

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Transforms

- Aggregate
- Filter Columns
- Filter Columns Details
- Filter Rows
- Join
- JSON Query

Text

Models

- Anomaly Detection
- Association
- Classification
- Clustering
- Feature Extraction
- Model
- Model Details
- Regression

Predictive Queries

- Evaluate and Apply
- Linking Nodes

Oracle Data Miner - Edit Filter Columns Node

☒ Show Attribute Importance
☒ Show Data Quality

Columns

Name	Type	Output	Rank	Importance	% Null	% Unique	% Constant	Hints
XYZ BANK_FUNDS	NUMBER	→	1	0.2039	0	21.3511	35.6016	
XYZ N_TRANS_ATM	NUMBER	→	2	0.1217	0	0.4438	21.499	
XYZ N_TRANS_TELLER	NUMBER	→	3	0.1213	0	0.4931	32.3471	
XYZ MONEY_MONTHLY_OVERDRAWN	NUMBER	→	4	0.1207	0	18.787	16.3215	
XYZ T_AMOUNT_AUTOM_PAYMENTS	NUMBER	→	5	0.1054	0	59.9606	21.5483	
XYZ MONTHLY_CHECKS_WRITTEN	NUMBER	→	6	0.0848	0	0.9369	19.428	
XYZ N_OF_DEPENDENTS	NUMBER	→	7	0.0315	0	0.3452	34.4181	
XYZ TIME_AS_CUSTOMER	NUMBER	→	8	0.0221	0	0.2465	31.3116	
XYZ CHECKING_AMOUNT	NUMBER	→	9	0.0183	0	30.3254	62.7712	
XYZ CREDIT_BALANCE	NUMBER	→	10	0.0135	0	9.4181	90.5819	
XYZ N_TRANS_KIOSK	NUMBER	→	11	0.0084	0	0.5424	43.6884	
XYZ MORTGAGE_AMOUNT	NUMBER	→	12	0.0073	0	21.2525	23.1262	
XYZ SEX	VARCHAR2	→	13	0.0064	0	0.0986	66.075	
XYZ MARITAL_STATUS	VARCHAR2	→	14	0.0056	0	0.2465	34.5168	
XYZ HOUSE_OWNERSHIP	NUMBER	→	15	0.004	0	0.1479	71.1538	

Help OK Cancel

Thumbnail:

BUY_INSURANCE_WORKFLOW - ...

- Clustering Segmentation
- Explore/Profile Data
- Explore Data
- Prediction Queries by Region
- Prediction Query
- Filter Cols_Attrib Importance
- Predictive Models for Student
- Class Build
- Predictive models ind TEXT
- Filter Columns Details
- Scatter Box plots etc.
- Graph
- Graph 1
- Anomaly Detection Query
- CUST_INSUR_LTV
- CUST_INSUR_LTV_APPLY
- MINING_DATA_TEXT_BUILD_V
- CUST_INSUR_LTV1
- Model Details

Components: Workflow Editor, Data

Data Source, Explore Data, SQL Query, Update Table

Filter Columns, Filter Columns Details, Join, JSON Query

Association, Classification, Feature Extraction, Model, Regression

Predictive Queries, Evaluate and Apply, Linking Nodes

dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData Miner

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY_INSURANCE_WORKFLOW

CARS_DATA_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

Thumbnail

BUY_INSURANCE_WORKFLOW - ...

Reports

Clustering Segmentation

Explore/Profile Data

Explore Data

Prediction Queries by Region

Prediction Query

Filter Cols_Attr Importance

Filter Columns

Predictive Models for Student

Class Build

Predictive models ind TEXT

Filter Columns Details

Scatter Box plots etc.

Graph

Graph 1

Anomaly Detection Query

CUST_INSUR_LTV

CUST_INSUR_LTV_APPLY

MINING_DATA_LTV_BUILD_V

CUST_INSUR_LTV1

Start Page

BUY_INSURANCE_WORKFLOW

CUST_INSUR_LTV1

Graph 1

100%

Parallel Query Off

Graph 1

CUST_INSUR_LTV1

Explore Data

Filter Columns

Class Build

Connect

Run

Force Run

Edit...

Advanced Settings...

View Models

View Test Results

Compare Test Results

Generate Apply Chain

Show Event Log

Validate Parents

Deploy

Cut

Copy

Paste

Extended Paste...

Select All

Parallel Query ...

Copy Image to Clipboard

Save Image As...

Go to Properties

Navigate

Class Build 1 - Properties

Find

Models

Build

Test

Details

Model Settings

Name

Output

Build

Test

CLAS_GLM_5_2

CLAS_SVM_5_2

CLAS_DT_4_2

CLAS_NB_4_2

9/14/15 5:17 PM

9/14/15 5:17 PM

9/14/15 5:17 PM

9/14/15 5:17 PM

Automatic

Automatic

Automatic

Automatic

Generalized Linear Model

Support Vector Machine

Decision Tree

Naive Bayes

Comment

Components

Workflow Editor

Data

Create Table or View

Data Source

Explore Data

Graph

SQL Query

Update Table

Transforms

Aggregate

Filter Columns

Filter Columns Details

Filter Rows

Join

JSON Query

Text

Models

Anomaly Detection

Association

Classification

Clustering

Feature Extraction

Model

Model Details

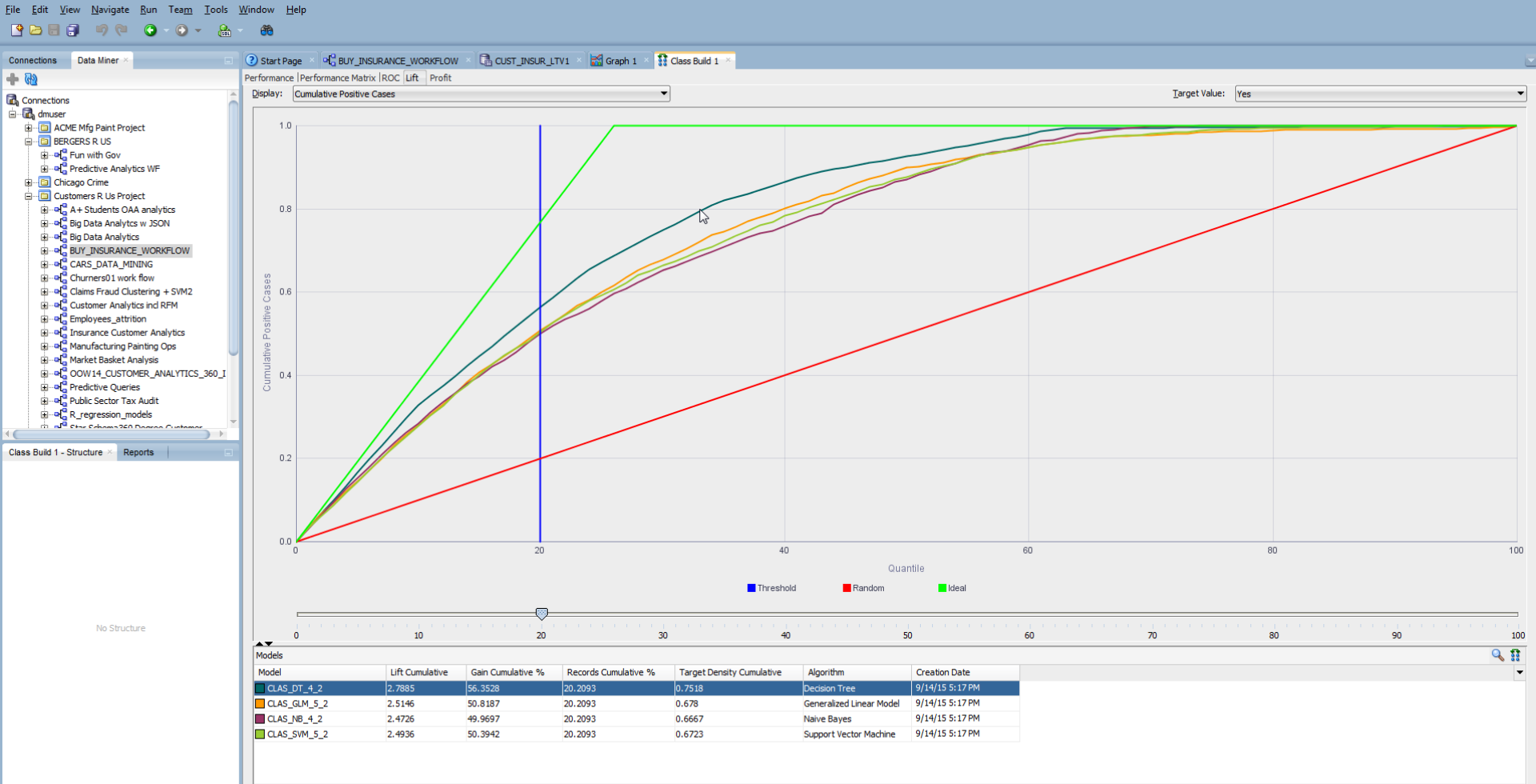
Regression

Predictive Queries

Evaluate and Apply

Linking Nodes

Synchronization complete



dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

FileEditViewNavigateRunDiagramTeamToolsWindowHelp

ConnectionsData Miner

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY_INSURANCE_WORKFLOW

CARS_DATA_MINING

Churners01 work flow

Claims Fraud Clustering + SVM2

Customer Analytics ind RFM

Employees_attrition

Insurance Customer Analytics

Manufacturing Painting Ops

Market Basket Analysis

Thumbnail

CLAS_DT_4_2 - StructureReports

Node: 0

Node: 1

Node: 15

Node: 16

Node: 2

Node: 3

Node: 4

Node: 5

Node: 17

Node: 18

Node: 19

Node: 7

Node: 8

Node: 20

Node: 21

Node: 9

Node: 22

Node: 23

Node: 10

Start PageBUY_INSURANCE_WORKFLOWCUST_INSUR_LTV1Graph 1CLAS_DT_4_2

TreeSettings

100%

Maximum Target Values: 2

Save Rules...

Confidence: 86.55%

No : 1319 (86.55%)

Yes : 205 (13.45%)

Confidence: 72.46%

No : 121 (72.46%)

Yes : 46 (27.54%)

Confidence: 72.08%

No : 615 (27.92%)

Yes : 1588 (72.08%)

Split: MONEY_MONTHLY_OVERDRAWN

Confidence: 64.15%

No : 823 (64.15%)

Yes : 460 (35.85%)

Split: N_TRANS_ATM

Confidence: 82%

No : 123 (82.00%)

Yes : 27 (18.00%)

Node: 9

Prediction: Yes

Support: 873 (9.46%)

Confidence: 86.25%

No : 120 (13.75%)

Yes : 753 (86.25%)

Split: T_AMOUNT_AUTOM_PAYMENTS

Node: 8

Prediction: Yes

Support: 1330 (14.42%)

Confidence: 62.78%

No : 495 (37.22%)

Yes : 835 (62.78%)

Split: T_AMOUNT_AUTOM_PAYMENTS

Node: 6

Prediction: No

Support: 1218 (13.2%)

Confidence: 62.4%

No : 760 (62.40%)

Yes : 458 (37.60%)

Split: T_AMOUNT_AUTOM_PAYMENTS

Node: 17

Prediction: No

Support: 65 (0.7%)

Confidence: 96.92%

No : 63 (96.92%)

Yes : 2 (3.08%)

Node: 22

Prediction: Yes

Support: 851 (9.22%)

Confidence: 88.25%

No : 100 (11.75%)

Yes : 791 (88.25%)

Node: 23

Prediction: No

Support: 22 (0.24%)

Confidence: 90.91%

No : 20 (90.91%)

Yes : 2 (9.09%)

Node: 20

Prediction: Yes

Support: 1244 (13.48%)

Confidence: 66%

No : 423 (34.00%)

Yes : 821 (66.00%)

Node: 21

Prediction: No

Support: 86 (0.93%)

Confidence: 83.72%

No : 72 (83.72%)

Yes : 14 (16.28%)

Node: 19

Prediction: No

Support: 418 (4.53%)

Confidence: 74.88%

No : 313 (74.88%)

Yes : 105 (25.12%)

Node: 18

Prediction: No

Support: 800 (8.67%)

Confidence: 55.88%

No : 447 (55.88%)

Yes : 353 (44.12%)

RuleSurrogatesTarget Values

Node Rule:

If BANK_FUNDS > 246

And CHECKING_AMOUNT <= 282

And CREDIT_BALANCE <= 2445

And MONEY_MONTHLY_OVERDRAWN > 54.095

And T_AMOUNT_AUTOM_PAYMENTS <= 14993

Then Yes

Confidence	0.882491186839013
Support	0.0922393236505279

Components

No Components

dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

ORACLE

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

62

File Edit View Navigate Run Diagram Team Tools Window Help

Connections Data Miner Start Page BUY_INSURANCE_WORKFLOW CUST_INSR_LTV1 Graph 1 CLAS_DT_4_2

Connections

- dmuser
- ACME Mfg Paint Project
- BERGERS R US
 - Fun with Gov
 - Predictive Analytics WF
 - Chicago Crime
- Customers R Us Project
 - A+ Students OAA analytics
 - Big Data Analytics w JSON
 - Big Data Analytics
 - BUY_INSURANCE_WORKFLOW
 - CARS_DATA_MINING
 - Churners01 work flow
 - Claims Fraud Clustering + SVM2
 - Customer Analytics ind RFM
 - Employees_attrition
 - Insurance Customer Analytics
 - Manufacturing Painting Ops

Thumbnail

BUY_INSURANCE_WORKFLOW - ... Reports

- Clustering Segmentation
- Explore/Profile Data
- Explore Data
- Prediction Queries by Region
- Prediction Query
- Filter Cols_Attrib Importance
- Filter Columns
- Predictive Models for Student
- Class Build
- Predictive models ind TEXT
- Class Build 1
- Filter Columns Details
- Scatter Box plots etc.
- Graph
- Graph 1
- Anomaly Detection Query
- CUST_INSR_LTV
- CUST_INSR_LTV_APPLY
- MINING_DATA_TEXT_BUILD_V
- CUST_INSR_LTV1

Workflow Diagram:

Graph 1

CUST_INSR_LTV1

Explore Data

Filter Columns

Class Build 1

CUST_INSR_LTV_APPLY1

Context Menu:

- Connect
- Run
- Force Run
- Edit...
- View Data
- Generate Apply Chain
- Show Event Log
- Validate Parents
- Deploy
- Save SQL
- Cut
- Copy
- Paste
- Extended Paste...
- Select All
- Parallel Query ...
- Copy Image to Clipboard
- Save Image As...
- Go to Properties
- Navigate

Apply - Properties

Predictions

Additional Output

Cache

Details

Output Apply Columns

Column	Function	Parameter(s)	Model	Node
CLAS_NB_4_2_PROB_Yes	Prediction Probability	Prediction: Yes	CLAS_NB_4_2	Class Build 1
CLAS_NB_4_2_PDET	Prediction Details	Prediction: Yes, Sort: Absolute, Length: 5	CLAS_NB_4_2	Class Build 1

Components

Workflow Editor

Data

- Create Table or View
- Data Source
- Explore Data
- Graph
- SQL Query
- Update Table

Transforms

- Aggregate
- Filter Columns
- Filter Columns

Text

Models

- Anomaly Detection
- Association
- Classification
- Clustering
- Feature Extraction
- Model

Predictive Queries

Evaluate and Apply

- Apply
- Test

Linking Nodes

dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

FileEditViewNavigateRunTeamToolsWindowHelp

ConnectionsData Miner

Start PageBUY_INSURANCE_WORKFLOWCUST_INSR_LTV1Graph 1CLAS_DT_4_2Apply

DataColumnsSQL

ViewActual DataSort...Parallel Query Off...FilterEnter Where Clause

CLAS_NB_4_2_PROB_Yes	CLAS_NB_4_2_POET	CUST_ID	LAST	N_MORTGAGES	SALARY
1 0.9998904466629028	<Details algorithm="Naive Bayes" clas...	CU3367	LOUISE	1	63,966
2 0.9998904466629028	<Details algorithm="Naive Bayes" clas...	CU9432	CLEOTILDE	1	96,573
3 0.9998646974563599	<Details algorithm="Naive Bayes" clas...	CU6274	SUDIE	1	69,159
4 0.9998481869697571	<Details algorithm="Naive Bayes" clas...	CU2126	DWANA	1	65,581
5 0.9998325705528259	<Details algorithm="Naive Bayes" clas...	CU12392	FLOSSIE	1	93,595
6 0.999816358089447	<Details algorithm="Naive Bayes" clas...	CU13307	JULIO	1	70,535

Components

No Components

Connections

dmuser

ACME Mfg Paint Project

BERGERS R US

Fun with Gov

Predictive Analytics WF

Chicago Crime

Customers R Us Project

A+ Students OAA analytics

Big Data Analytics w JSON

Big Data Analytics

BUY_INSURANCE_WORKFLOW

CARS_DATA_MINING

Churners01 work flow

Claims Fraud Clustering

Customer Analytics incl RF

Employees_attrition

Insurance Customer Analy

Manufacturing Painting Op

Market Basket Analysis

OOW14_CUSTOMR_ANA

Predictive Queries

Public Sector Tax Audit

R_regression models

Apply - Structure

Reports

No Structure

View Value

Find

```
<Details algorithm="Naive Bayes" class="Yes">
  <Attribute name="LTV_BIN" actualValue="MEDIUM" operator="in" range="LOW,MEDIUM,VERY HIGH" weight=".987" rank="1"/>
  <Attribute name="N_TRANS_WEB_BANK" actualValue="2300" operator="greaterThan" value="1419.5" weight=".963" rank="2"/>
  <Attribute name="LTV" actualValue="21389.75" operator="greaterThan" value="6861.625" weight=".848" rank="3"/>
  <Attribute name="CAR_OWNERSHIP" actualValue="1" weight=".813" rank="4"/>
  <Attribute name="AGE" actualValue="51" operator="between" range="(17.5:55.5]" weight=".8" rank="5"/>
</Details>
```

Close

33 0.9997174739837646	<Details algorithm="Naive Bayes" clas...	CU7142	RAELENE	1	86,368
34 0.9997147917747498	<Details algorithm="Naive Bayes" clas...	CU1717	MARGARETA	1	58,472
35 0.9997147917747498	<Details algorithm="Naive Bayes" clas...	CU2659	TIMIKA	1	68,203
36 0.9997147917747498	<Details algorithm="Naive Bayes" clas...	CU660	JINNY	1	61,370
37 0.9997106790542603	<Details algorithm="Naive Bayes" clas...	CU737	KINA	1	68,304
38 0.9997106790542603	<Details algorithm="Naive Bayes" clas...	CU3016	LATRINA	1	66,364
39 0.9997106790542603	<Details algorithm="Naive Bayes" clas...	CU3846	GWENDA	1	71,093
40 0.9996994733810425	<Details algorithm="Naive Bayes" clas...	CU12872	SHARRI	1	75,125
41 0.99969827162835022	<Details algorithm="Naive Bayes" clas...	CU616	YOLANDA	1	63,585

dmuser/Customers R Us Project/BUY_INSURANCE_WORKFLOW

ORACLE

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

64

Connections

Data Miner

- Customers R Us Project
 - A+ Students OAA analytics
 - Big Data Analytics w JSON
 - Big Data Analytics
 - BUY_INSURANCE_WORKFLOW
 - CARS_DATA_MINING
 - Churners01 work flow
 - Claims Fraud Clustering + SVM2
 - Customer Analytics ind RFM
 - Employees_attrition**
 - Insurance Customer Analytics
 - Manufacturing Painting Ops
 - Market Basket Analysis
 - OOW14_CUSTOMER_ANALYTICS_360
 - PAW Cust Analytics
 - Predictive Queries
 - Public Sector Tax Audit
 - R_regression_models
 - RFM Customer Analytics
 - RFM CUSTOMER ANALYTICS2
 - Star Schema360 Degree Customer workflow

Thumbnail

Start Page

Employees_attrition

Hist. Employee data

CLAS_DT_1_10

Tree Settings

64%

Maximum Target Values: 2

Save Rules...

At Risk Employees

Data Columns | SQL

View: Cache Data | Sort... | Parallel Query Off... | Filter: Enter Where Clause

	CLAS_GLM_1_10_PROB_NO	CLAS_GLM_1_10_PDET
1	0.4790120104773802	<Details algorithm="Generalized Linea...
2	0.03718578019711771	<Details algorithm="Generalized Linea...
3	0.1457710019211804	<Details algorithm="Generalized Linea...
4	0.12768928733510213	<Details algorithm="Generalized Linea...
5	0.5698507390589098	<Details algorithm="Generalized Linea...
6	0.8124155905506835	<Details algorithm="Generalized Linea...
7	0.2038408741589256	<Details algorithm="Generalized Linea...
8	0.8968650734058157	<Details algorithm="Generalized Linea...
9	0.6750795593148387	<Details algorithm="Generalized Linea...
10	0.791574109722724	<Details algorithm="Generalized Linea...
11	0.000000587337939665	<Details algorithm="Generalized Linea...
12	0.148971355681	<Details algorithm="Generalized Linea...
13	0.5366586052554	<Details algorithm="Generalized Linea...
14	0.07470752471	<Details algorithm="Generalized Linea...
15	0.695885974645	<Details algorithm="Generalized Linea...
16	0.31135950403	<Details algorithm="Generalized Linea...
17	0.6143770649	<Details algorithm="Generalized Linea...
18	0.46131803526	<Details algorithm="Generalized Linea...
19	0.71284917635	<Details algorithm="Generalized Linea...
20	0.58226130772	<Details algorithm="Generalized Linea...
21	0.36108362457662586	<Details algorithm="Generalized Linea...
22	0.2988898709403637	<Details algorithm="Generalized Linea...
23	0.31307063556114767	<Details algorithm="Generalized Linea...

Rule

Surrogate

Node Rule:

If BONTU:

And HEAL:

And STOCI

And 2.5

And AMT_

Then NO

Confidence	0.6183533447684391
Support	0.15131066701271736

Copyright © 2015, 2016 Oracle and/or its affiliates. All rights reserved.

Find

```

Model" class="NO">
actualValue="5" weight="-.246" rank="1"/>
value="1" weight="-.046" rank="2"/>
e="F" weight=".036" rank="3"/>
actualValue="25" weight="-.022" rank="4"/>
27" weight="-.021" rank="5"/>

```

Close



Search



- CUST_INSUR_PROFILES
 - PROFILE_ID
 - PREDICTION
 - PREDICTION_COUNT
 - RECORD_COUNT
 - TOTAL_RECORD_COUNT
 - IS_LEAF_NODE
 - CONFIDENCE
 - SUPPORT
 - FULL_SIMPLE_RULE**
- My Calculations
 - Value Labels



Untitled

+ Click here or drag data elements to add a filter



Pie

Trellis Columns

Trellis Rows

Values (Slice)

PREDICTION_CO...

Category

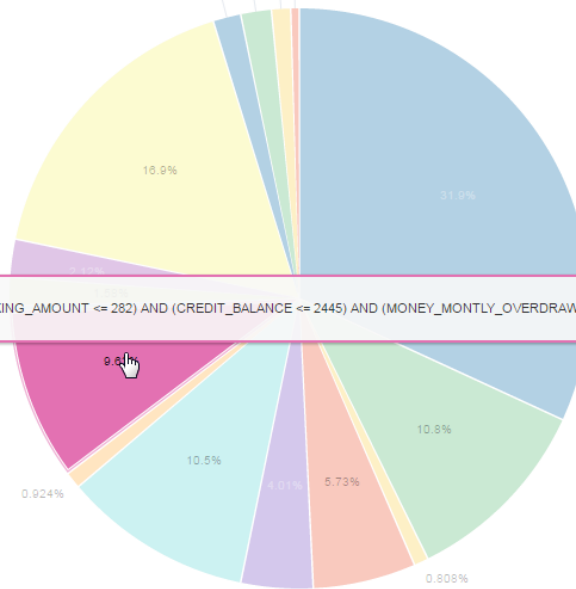
PREDICTION

Color

FULL_SIMPLE_RU..

PREDICTION_COUNT by PREDICTION, FULL_SIMPLE_RULE

0.475%
1.00%
1.71%
1.55%



PREDICTION_COUNT 751.00
FULL_SIMPLE_RULE (BANK_FUNDS > 246) AND (CHECKING_AMOUNT <= 282) AND (CREDIT_BALANCE <= 2445) AND (MONEY_MONTHLY_OVERDRAWN > 54.095) AND (T_AMOUNT_AUTOM_PAYMENTS <= 14993)
PREDICTION Yes

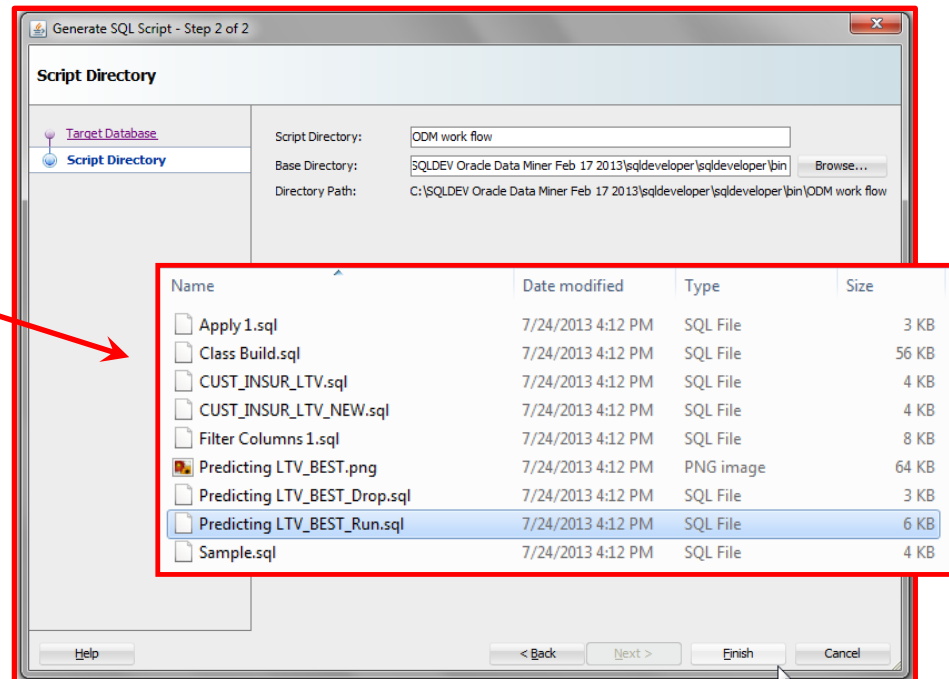
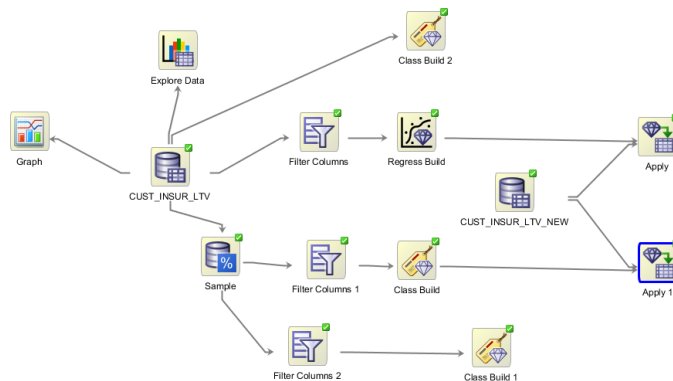
FULL_SIMPLE_RULE

Sharing, Automation and Deployment

Immediately Go to “Productionization” of Analytical Methodologies



- Share ODMr workflows
- Workflow API for 100% automation
 - Immediate deployment of data analyst's methodologies
- SQL Script Generation
 - Deploy methodology as SQL scripts



Fraud Prediction Demo

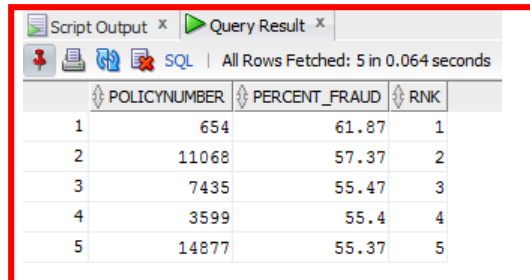
Automated In-DB Analytical Methodology



```
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;
```



	POLICYNUMBER	PERCENT_FRAUD	RNK
1	654	61.87	1
2	11068	57.37	2
3	7435	55.47	3
4	3599	55.4	4
5	14877	55.37	5

Automated Monthly “Application”!

Just add:

Create

View CLAIMS2_30

As

Select * from CLAIMS2

Where mydate > SYSDATE – 30

Time measure: set timing on;

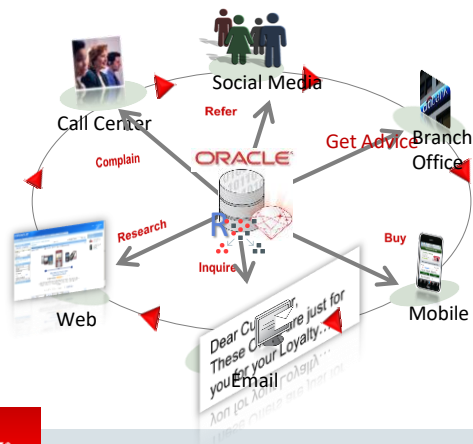
Oracle Advanced Analytics

Real-Time Scoring, Predictions and Recommendations



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

```
Select prediction_probability(CLAS_DT_1_64, 'Yes'  
    USING 7800 as bank_funds, 125 as checking_amount, 20 as  
    credit_balance, 55 as age, 'Married' as marital_status,  
    250 as MONEY_MONTHLY_OVERDRAWN, 1 as house_ownership)  
from dual;
```



Likelihood to respond:

Query Result	
All Rows Fetched: 1 in 0 seconds	
PREDICTION_PROB...	0.8382936507936...

OAA/Oracle Data Miner 4.1 HOL

Uses Oracle by Example Free Online Tutorials

- There are 6 Tutorials

- The first tutorial is already done for you

- Recommend doing 3-5 Tutorials

1. [Using Oracle Data Miner 4.1](#)

2. [Star Schema Mining Using Oracle Data Miner 4.1](#)

3. [Text Mining with an EM Clustering Model Using Data Miner 4.1](#)

4. **Anomaly Detection (CLAIMS)** *See Instructor for assistance*

5. **Market Basket Analysis (SH.SALES)** *See Instructor for assistance*

Content List (click links below to view content)

08-SEP-2015
15 mins
★★★★★

[Setting Up Oracle Data Miner 4.1](#)

This tutorial covers the process of setting up Oracle Data Miner for use within Oracle SQL Developer 4.1.

08-SEP-2015
45 mins
★★★★★

[Using Oracle Data Miner 4.1](#)

This tutorial covers the use of Oracle Data Miner 4.1 to perform data mining against Oracle Database 12c. In this lesson, you examine the Oracle Data Miner GUI. The Oracle Data Miner GUI is included as an extension of Oracle SQL Developer, version 4.1.

08-SEP-2015
30 mins
★★★★★

[Star Schema Mining Using Oracle Data Miner 4.1](#)

This tutorial covers the use of Oracle Data Miner 4.1 to perform star schema mining activities against Oracle Database 12c Release 2.

08-SEP-2015
30 mins
★★★★★

[Text Mining with an EM Clustering Model Using Data Miner 4.1](#)

In this lesson, you learn how to use the EM algorithm in a clustering model while leveraging text mining enhancements that are included in Oracle Database 12c Release 2.

08-SEP-2015
30 mins
★★★★★

[Using Logistic Regression Models \(GLM\) to Predict Customer Affinity](#)

This tutorial covers the use of Oracle Data Miner 4.1 to leverage enhancements to the Oracle implementation of Generalized Linear Models (GLM) to predict customer affinity.

08-SEP-2015
30 mins
★★★★★

[Using the SQL Query Node With Oracle Data Miner 4.1](#)

This tutorial covers the use of the new SQL Query Node in an Oracle Data Miner 4.1 workflow.

08-SEP-2015
30 mins
★★★★★

[Using Predictive Queries With Oracle Data Miner 4.1](#)

This tutorial covers the use of Predictive Queries against mining data by Oracle Data Miner 4.1.

08-SEP-2015
30 mins
★★★★★

[Mining JSON Data Using Oracle Data Miner 4.1](#)

This tutorial covers the use of the JSON Query Node in an Oracle Data Miner 4.1 workflow in order to mine this Big Data format.



ORACLE®