OpenWorld 2018

Developing Predictive Applications with Oracle’s Machine Learning [DEV5758]

Move the Algorithms; Not the Data!

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

Kiran Bellare, Sr. Director Product Management, PaaS Cloud, Oracle
Developing Predictive Applications with Oracle’s Machine Learning [DEV5758]

Most data science projects don’t get beyond the data scientist and rarely “operationalize” their predictive models. By “moving algorithms to the data,” Oracle now embeds prebuilt analytical machine learning methodologies into applications. This presentation shows how it is done on premises or in the cloud and provides an inventory of current predictive applications, including HCM cloud, sales cloud, retail GBU and FinServ GBU applications, industry data models, and adaptive intelligence applications for manufacturing. Finally, it presents a preview of some new predictive applications in the pipeline. Come away with an understanding of how you too can build machine learning in to your applications.

**Code One Tracks:** Database, Big Data, and Data Science  
**Session Type:** Developer Session

**SPEAKERS**  
**Charlie Berger**, Sr. Director Product Management, Machine Learning, AI and Cognitive Analytics, Oracle  

- Monday, Oct 22, 12:30 p.m. - 1:15 p.m. | Moscone West - Room 2003
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.
Market Observations:

• Machine learning, predictive analytics & “AI” now *must-have* requirements
• Separate islands for data management and data science just don’t work
• Enterprises whose data science teams most rapidly extract insights and predictions win

Conclusions:

• Must “operationalize” ML insights and predictions throughout enterprise
• Multilingual Machine Learning: SQL, R, Python, Workflow UI, Notebooks, Embed ML in Apps
• Evolving towards *combined* Data Management + Machine Learning environment that can essentially to manage and “think” about data
Oracle’s Data Management and Machine Learning Architectural Strategy

• In-Database proprietary implementations of machine learning algorithms
• Leverage strengths of the Database and adds new ML tech
  – Counting, conditional probabilities, sort, rank, partition, group-by, collections, etc.
  – Parallel execution, bitmap indexes, partitioning, aggregations, recursion w/in parallel infrastructure, IEEE float, frequent itemsets, Automatic Data Preparation (ADP), Text processing, etc.
• Focus on intelligent ML defaults, simplification & automation to enable applications
  – ADP, xforms, binning, missing values, Prediction_Details, Predictive_Queries, Model_views
• Machine learning models built via PL/SQL script; scored via SQL functions (1st class DB objects)
  – “Smart scan” ML model scoring “push down”; Supports OLTP and ATPC environments
• Machine Learning and Advanced Analytics are peer to rest of Oracle Data Mgmt features
  – Best ML & analytical development and deployment platform

```sql
select cust_id
from customers
where region = ‘US’
  and prediction_probability(churnmod, ‘Y’ using *) > 0.8;
```

True power evident when scoring models using SQL functions, e.g.
Oracle’s Machine Learning & Advanced Analytics
Fastest Way to Deliver Enterprise-wide Predictive Analytics

Major Benefits

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

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

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

**ANOMALY DETECTION**
- One-Class SVM

**TIME SERIES**
- State of the art forecasting using Exponential Smoothing.
- Includes all popular models e.g. Holt-Winters with trends, seasons, irregularity, missing data

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

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

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

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

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

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

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

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

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

**EXPORTABLE ML MODELS**
- REST APIs for deployment

---

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

Copyright © 2018, Oracle and/or its affiliates. All rights reserved.
Multiple Data Scientist User Roles Supported

Oracle’s Machine Learning/Advanced Analytics

DBAs

Application Developers

R, Python Users, Data Scientists

OML Notebook Users

Data Analysts, Citizen Data Scientists
OAA Model Build and Real-time SQL Apply Prediction

Simple SQL Syntax

ML Model Build (PL/SQL)

```sql
begin
    dbms_data_mining.create_model('BUY_INSUR1', 'CLASSIFICATION', 
        'CUST_INSUR_LTV', 'CUST_ID', 'BUY_INSURANCE', CUST_INSUR_LTV_SET);
end;
/
```

Model Apply (SQL query)

```sql
Select prediction_probability(BUY_INSUR1, 'Yes'
    USING 3500 as bank_funds, 825 as checking_amount, 400 as credit_balance, 22 as age, 'Married' as marital_status, 93 as MONEY_MONTLY_OVERDRAWN, 1 as house_ownership)
from dual;
```

| PREDICTION_PROBABILITY(BUY_INSUR1,YESSUSING3500ASBANK_FUNDS,825ASCHECKING_AMOUNT,400ASCREDIT_BALANCE | 0.9276956709910801 |
## OAA Algorithm Performance

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Time (sec.)</th>
<th>Number of Attributes</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM SGD Solver</td>
<td>83s</td>
<td>900</td>
</tr>
<tr>
<td>GLM Classification</td>
<td>180s</td>
<td>9000</td>
</tr>
<tr>
<td>SVM IPM Solver</td>
<td>811s</td>
<td>900</td>
</tr>
<tr>
<td>GLM Regression</td>
<td>59s</td>
<td>900</td>
</tr>
<tr>
<td>K-Means</td>
<td>168s</td>
<td>9000</td>
</tr>
</tbody>
</table>

- In-memory 640 million rows, Airline On-Time dataset
- SPARC M8-2 hardware
- Courtesy of the SPARC Performance Team
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
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

Machine Learning Algorithms

Data + Domain Knowledge

ML Algorithms

Models, Insights, Predictions
How long does it take to put a defined model into operational use?

Length of time to put a model into production.
Based on 141 respondents who stated they are doing this today
Oracle Applications that embed Oracle’s Machine Learning Algorithms
Oracle Machine Learning Applications: GA

**HCM Cloud**
Workforce Predictions

**CRM Sales Cloud**
Sales Prediction

**Retail GBU**
Customer Insights, Customer Segmentation

**Adaptive Intelligent Appl for Manufacturing**

**CPQ Cloud**
Configure, Price, Quote

**Industry Data Models**
Comms, SNA, Utilities, Airlines, Retail, ...

**Oracle E-Business Suite**
Spend Classification

**Oracle Identity Mgmt**
Adaptive Access Mgmt

**FSGBU**
Analytical Applications Infrastructure
Oracle Machine Learning Applications: GA

HCM Cloud
Workforce Predictions

CRM Sales Cloud
Sales Prediction

Retail GBU
Customer Insights,
Customer Segmentation

Adaptive Intelligent Appl
for Manufacturing

CPQ Cloud
Configure, Price, Quote

Industry Data Models
Comms, SNA, Utilities, Airlines, Retail, ...

Oracle E-Business Suite
Spend Classification

Oracle Identity Mgmt
Adaptive Access Mgmt

FSGBU
Analytical Applications
Infrastructure
Application Store-Level Customer Segmentation

Retail Store Assortments Based on Local Customer Preferences

Machine Learning Use Cases

• Optimize store product assortments
• Coordinate marketing and merchandising
• Guide 1-to-1 promotions

Reasons why using

• Eliminate data movement and scale to 50 million Sku/store combinations
• One version of truth from central data warehouse
• Security

Future ML Plans

• Potential integration with ODC
Market Basket and Demand Transference Insights

Demand Transference Analysis

**Appl’s ML Use Cases**

- Gain actionable insights into customer behavior
- Map product attribute affinities to customer attributes
- Optimize shelf assortment breadth and depth

**Reasons why using/like OAA’s ML in Appl**

- Easy citizen scientist user workflow
- Algorithmic extensibility
- Security

**Future ML Plans**

- Tighter linkage to physical shelf and store constraints
Recent Oracle Applications that embed Oracle’s Machine Learning Algorithms
Oracle Adaptive Intelligent (AI) Apps for Manufacturing
Achieve Manufacturing Operational Excellence using Machine Learning & AI

ML Use Cases
• Insights (Patterns and Correlations Analysis)
  – Discover key influencers and patterns that affect yield & quality
• Predictive Analytics
  – Predictive critical outcomes during manufacturing to minimize losses

Reasons why using/like OAA’s ML
• Easy-to-integrate R & PL/SQL APIs for many ML algorithms
• In-database execution & scalable performance
• Enterprise grade support for OAA ML

Future ML Plans
• Feature Significance using Attribute Importance of OAA
• Unsupervised Learning (Clustering) on manufacturing big data
• Image, audio, and video analysis for product defects
76% of Strawberry Jam lots that Failed Consistency Test from Jan 1 – June 30 had the following Factors:

- Operator Skill was Level 2 @ Mixing
- Production was during 2nd Shift
- Pectin (Ingredient) Quantity was > 5.5 LB
- Supplier of Strawberry was "Berry Farms"
- Blender Max. Speed was 590 RPM in 0 -15 min
- External Humidity was > 72 % @ Mixing Work Center
Configure, Price, Quote (CPQ)
Accelerate Sales, Improve Margins, and Control Sales Operations

Functionality
• CPQ Cloud introduced machine learning Price Optimization for Transactions
  • Predicts the likelihood that a customer will buy at the current level of pricing of a Transaction
  • Win Probability feature to guide sales users in their negotiation of pricing and discounts

Advantages of Using OAA
• In-Database;
  • Ease of development, integration and deployment

Timetable and Road Map
• GA December 2017
E-Business Suite: Depot Repair
Repair Faster and Cheaper by Predicting Root Cause and Best Fix

Functionality
• Predict root cause
• Predict best fix
• Automatically apply accepted recommendations

Advantages of Using OAA
• Ease of integration & deployment
• Strength of support network
• The right algorithms

Timetable and Road Map
• GA now: 12.2.7 Patch #28263445
• Future: text mine technician notes
• Future: warranty fraud detection
• Future: greater transparency
Let’s Take a Quick Look at Oracle Human Capital Management’s (HCM Cloud) Workforce Predictions as an Embedded Example
Enabling Predictive Enterprise Applications

HCM Predictive Workforce

• Integrated data management + embedded predictive analytics
• Full 360 degree employee view
• Single source of HCM data data
• Interactive dashboards and “What if” analysis
• Customizable if desired to add input variables to predictive models
• Mobile + Oracle Cloud solutions

Link to HCM Predictive Workforce demo
HCM Predictive Workforce

Predict Employee Attrition, Performance and What If? Analysis

• Identify employees’ likelihood to leave and predicted performance
• Top reasons for expected behavior
• Interactive “What if?” analysis

Link to Oracle HCM on O.com
HCM Predictive Workforce demo
-- 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
--
-- List of attribute names ranked by their importance value.
-- The larger the value, the more impact that attribute has
-- on causing variation in the target column.
--
column attribute_name format a40
column explanatory_value format 9.999
SELECT attribute_name, explanatory_value, rank
FROM ai_explain_output
ORDER BY rank, attribute_name;
Given demographic data about a set of customers, predict the customer response to an affinity card program using a classifier based on Decision Trees algorithm.

-- CREATE A NEW MODEL
-- Build a DT model

BEGIN
  DBMS_DATA_MINING.CREATE_MODEL(
    model_name => 'DT_SH_Clas_sample',
    mining_function => dbms_data_mining.classification,
    data_table_name => 'mining_data_build_v',
    case_id_column_name => 'cust_id',
    target_column_name => 'affinity_card',
    settings_table_name => 'dt_sh_sample_settings');
END;
/

SELECT T.cust_id, S.prediction, S.probability, S.cost
FROM (SELECT cust_id,
          PREDICTION_SET(dt_sh_clas_sample COST MODEL USING *) pset
          FROM mining_data_apply_v
          WHERE cust_id < 100011) T,
          TABLE(T.pset) S
ORDER BY cust_id, S.prediction;
OAA Oracle Data Mining SQL Sample Programs
Starter SQL and PL/SQL Scripts for Learning and Fast-Starts

- **Oracle Database PL/SQL Packages and Types Reference** for syntax of the PL/SQL API
- **Oracle Data Mining Application Developer's Guide** is a Virtual Book with links for information on the use of the APIs
- **Oracle Database SQL Language Reference** for syntax of the SQL functions for model scoring
OAA Oracle Data Mining SQL Sample Programs

Directory Listing of the Data Mining Sample Programs

<table>
<thead>
<tr>
<th>dmaidemo.sql</th>
<th>dmsvrdeem.sql</th>
<th>dmshgrants.sql</th>
</tr>
</thead>
<tbody>
<tr>
<td>dmkmdemo.sql</td>
<td>dmdtxvlddemo.sql</td>
<td>dmglrdem.sql</td>
</tr>
<tr>
<td>dmsvddemo.sql</td>
<td>dmocdemo.sql</td>
<td>dmstardemo.sql</td>
</tr>
<tr>
<td>dmardemo.sql</td>
<td>dmtxtnmf.sql</td>
<td>dmhpdemo.sql</td>
</tr>
<tr>
<td>dmnbdemo.sql</td>
<td>dmemdemo.sql</td>
<td>dmsvcdem.sql</td>
</tr>
<tr>
<td>dmsvodem.sql</td>
<td>dmsh.sql</td>
<td></td>
</tr>
<tr>
<td>dmdtdemo.sql</td>
<td>dmtxtsvm.sql</td>
<td></td>
</tr>
<tr>
<td>dmndndemo.sql</td>
<td>dmglcdem.sql</td>
<td></td>
</tr>
</tbody>
</table>
New Oracle Machine Learning for Python — OML4Py
(Available Soon on OTN for Download as OAA Add-in Component — like Oracle R Enterprise)
Traditional Python and Database Interaction

- Access latency
- Paradigm shift: Python → SQL → Python
- Memory limitation – data size
- Ad hoc production deployment
- Issues for backup, recovery, security

Python script
cron job

mxODBC, pyodbc, turboodbc, JayDeBeApi, cx_Oracle
Oracle Machine Learning for Python

Oracle Advanced Analytics Option to 18c+

• Similar architecture to OAA’s Oracle R Enterprise
• OML4Py Transparency Layer
  – Use Oracle Database as High Performance Computing environment
• OML4Py OAA Model Build and Apply
  – Use OAA parallel and distributed ML algorithms
  – Manage Python scripts and Python objects in Oracle Database
• OML4Py Embedded Python
  – Make callout to Python packages
  – Integrate Python results into applications via SQL
OML4Py Transparency Layer

- Leverages proxy objects for database data:
  oml.DataFrame
  # DB table Boston
  DATA = core.sync(table = 'BOSTON')
  # Pandas DataFrame data
  DATA = core.create(data, table = 'BOSTON')

- Overloads Python functions translating functionality to SQL

- Use familiar Python syntax to manipulate database data

```python
DATA.shape
DATA.head()
DATA.describe()
DATA.std()
DATA.skew()
train_dat, test_dat = DATA.split()
train_dat.shape
test_dat.shape
```
New AutoML included initially available in Oracle Machine Learning for Python—OML4Py
(Available Soon on OTN for Download as OAA Add-in Component —like Oracle R Enterprise)

See AutoML demos at Research and Advanced Development at Oracle Labs demo pod – CSS06
Automated Machine Learning (AutoML): Motivation

- Many best candidate models for each prediction task, for example:
  - Classification: SVM, Random Forest, NNs, etc
- Hyper-parameters tuning for each model
  - Multiple iterations for each model training
- Explosion of features; which are relevant?
  - Feature Selection is a key pre-processing problem
- Laborious & time-consuming
- ML system defaults based on data good
  - But significant improvements are possible in several cases
Model Build & Real-time SQL Apply Prediction

OML4Py Syntax

ML Model Build Best Model (Python)

```python
ms = ModelSelection(
    mining_function = 'classification',
    score_metric = 'accuracy')
best_model = ms.select(X_train, y_train)
```

Model Apply (Python)

```python
y_pred = best_model.predict(X_test)
```
Automated Machine Learning (AutoML) Components

Auto Feature Selection
- AutoFS
  - >50% reduction in features

Auto Model Selection
- AutoMS
  - 4x faster than exhaustive search

Auto Tuning of Hyperparameters
- AutoTune
  - Upto 24% accuracy improvement

Machine Learning Pipeline

Training Data

Inference
AutoTune Example of Tuning Neural Network

```python
at = Autotune(
    mining_function = 'classification',
    score_metric = 'accuracy')

evals = at.run('nn', X_train, y_train)

mod = evals['best_model']

y_pred = mod.predict(X_test)
```
AutoTune: Evaluation for OAA Neural Network

Avg Improvement of 2.86%

8-22% improvement for several datasets

OAA Neural Network - Default vs Tuned Accuracy
New Oracle Machine Learning Microservices
(Available Now to Internal Oracle Application Teams)
New Oracle Machine Learning Microservices (Available Now to Internal Oracle Application Teams)

- Services for building, storing, and deploying Oracle Machine Learning models
- Cognitive Services
  - Image and Text
- Documented RESTful APIs for application integration
- Containers (Docker images) for portability
- Model repository for storing, versioning, and comparing ML models
- Kubernetes support for container management
OML Image Microservices

- Cognitive APIs for images: tagging, classification, face detection, age, gender, inappropriate content, image similarity...
  - Deep Learning models & Transfer Learning
OML Image Microservices

• Cognitive APIs for images: tagging, classification, face detection, age, gender, inappropriate content, image similarity...
  – Deep Learning models & Transfer Learning
OML Image Microservices

• Cognitive APIs for images: tagging, classification, face detection, age, gender, inappropriate content, image similarity...
  – Deep Learning models & Transfer Learning
OML Text Microservices

• Cognitive Text API
  – Summary
  – Key words
  – Topics
  – Sentiment
  – Similarity

• Uses Explicit Semantic Analysis (ESA) Wikipedia model
OML Text Microservices

- Cognitive Text API
  - Summary
  - Key words
  - Topics
  - Sentiment
  - Similarity

- Uses Explicit Semantic Analysis (ESA) Wikipedia model
OML Text Microservices

- Cognitive Text API
  - Summary
  - Key words
  - Topics
  - Sentiment
  - Similarity

- Uses Explicit Semantic Analysis (ESA) Wikipedia model
OML Text Microservices

• Cognitive Text API
  – Summary
  – Key words
  – Topics
  – Sentiment
  – Similarity

• Uses Explicit Semantic Analysis (ESA) Wikipedia model
Score Model

To score a row of data, enter some values in the value column of the attribute table.

Model: buyins_rf (CLASSIFICATION)

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>BANK_FUNDS</td>
<td>3200</td>
<td>NUMERICAL</td>
</tr>
<tr>
<td>MONEY_MONTHLY_OVERDRAWN</td>
<td>55</td>
<td>NUMERICAL</td>
</tr>
<tr>
<td>MONTHLY_CHECKS_WRITTEN</td>
<td></td>
<td>NUMERICAL</td>
</tr>
<tr>
<td>N_OF_DEPENDENTS</td>
<td></td>
<td>NUMERICAL</td>
</tr>
<tr>
<td>N_TRANS_ATM</td>
<td></td>
<td>NUMERICAL</td>
</tr>
<tr>
<td>N_TRANS_TELLER</td>
<td></td>
<td>NUMERICAL</td>
</tr>
<tr>
<td>T_AMOUNT_AUTOM_PAYMENTS</td>
<td></td>
<td>NUMERICAL</td>
</tr>
</tbody>
</table>

Score  Close

Results

Classification

<table>
<thead>
<tr>
<th>Label</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>0.5319246211305034</td>
</tr>
<tr>
<td>Yes</td>
<td>0.4680753788694968</td>
</tr>
</tbody>
</table>

Prediction Detail
/oaa-scoring/services/myservices/uri

GET

Gets the details for an OAA native endpoint. This consists of the model metadata along with other descriptive information.

DELETE

Removes a previously created endpoint identified by the unique uri parameter.
Content and Experience Cloud

PRO 4648
Deliver Personalized Experiences and Automate Content Management with AI and ML

Tuesday 10/23
5:45p
Marriott Marquis (Yerba Buena Level) Nob Hill C/D

Kiran Bellare
Content and Experience Cloud (CEC)

- Originally Documents Cloud Service which focused on EFSS (Enterprise File Sync and Share) for simple document management, focus was on: Files, Folders and Sharing

- Migrating toward Web Content Management and Customer Experience, focus shifting to: Digital Assets, Content Items, Web Sites, Multi-channel delivery, and the Customer Experience.
Content and Experience Cloud

• Upon instance delivery, CEC is an empty container
  • CEC has no specific target business focus
  • Customers leverage the toolset to construct their environments: content types, assets, delivery channels, web site pages, custom web components, etc.
  • Customers then develop specific content (images, descriptions, blogs, articles, summaries, reviews, etc)

• Phase 1: Content Development and Curation
  • Apply non-supervised learning algorithms, Wikipedia based models as common baseline especially for smaller repositories.
  • Text and Image keyword tagging
  • Identify inappropriate content
  • Auto-classification via similarity match against customer defined categories or hierarchies
  • Similar content via similarity recommendations
  • Sentiment analysis on visitor comments

• OML Microservices
  • Cognitive Image: Classification, NSFW
  • Cognitive Text: Topics, Keywords, Similarity, Sentiment
**Phase 2: Custom Models**
- Add product functionality (or leverage service) to train new image identification model
  - Submit sample images and keywords, train on samples, verify test cases, adjust, repeat
- Same for customer specific text domains: Legal, pharmaceutical, manufacturing, etc.

**Phase 3: Recommender System**
- Leverage Machine Learning models to drive site visitors to the next step in their “journey”
- Metadata from CEC repository as features
- Visitor Demographics from BlueKai
- Oracle Profile data from Master Customer Profile Service (MCPS)
- Web activity history from Oracle Infinity (Google Analytics, etc).
Analytics and Data Summit
All Analytics. All Data. No Nonsense.
March 12 – 14, 2019

Formerly called the BIWA Summit with the Spatial and Graph Summit
Same great technical content...new name!
www.AnalyticsandDataSummit.org