This session will be recorded

Oracle Machine Learning

AskTOM Office Hours – Feature Highlight

Batch scoring using OML Services on Autonomous Database

with Miles Novotny and Sherry LaMonica

Hosts: Mark Hornick

Product Management, Oracle Machine Learning

OML Services Batch Scoring on Autonomous Database

Speakers: Miles Novotny and Sherry LaMonica

OML Services on Oracle Autonomous Database now supports batch scoring for in-database models using REST endpoints. Users can build in-database models that reside in your database schema, where deployment is immediate via SQL queries for both batch and real-time scoring applications. With OML Services, those same models can now be deployed through REST endpoints hosted in Oracle Autonomous Database to enable batch scoring. This session includes a demonstration using a REST client with Oracle Autonomous Database.

Poll #1: Use of OML Services

Which features of OML Services have you used?

- Managing/deploying in-database models
- Managing/deploying ONNX-format models
- Data monitoring
- Model monitoring
- Singleton or small-batch scoring

How important of a feature do you consider asynchronous large-batch scoring for application developers?

ORACLE

Oracle Machine Learning

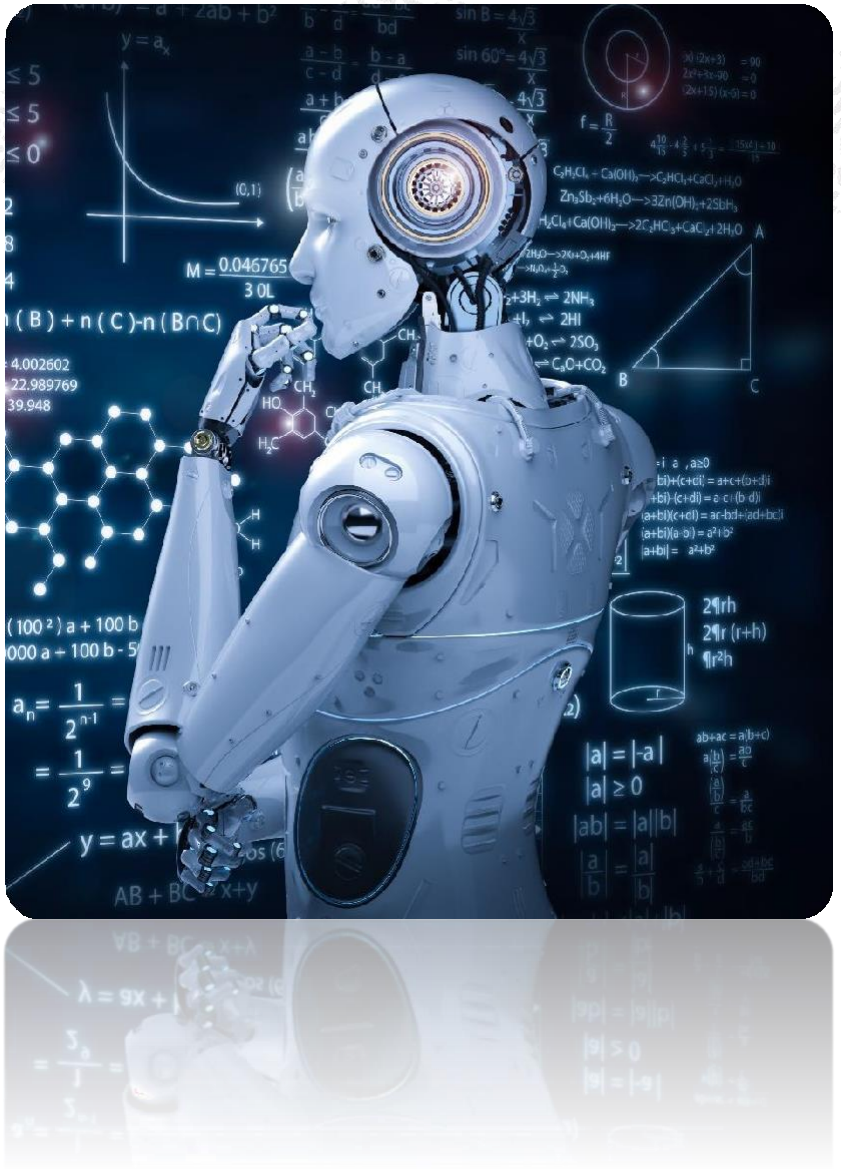
AskTOM Office Hours – Feature Highlight

Batch scoring using OML Services

Miles Novotny and Sherry LaMonica

Agenda

- OML Services Overview
- OML Services Batch Scoring
- Workflow
- Demonstration
- Q&A



OML Services

Supports model management, scoring, and monitoring using REST endpoints

- Available on Oracle Autonomous Database-Shared
- Enable key elements of overall enterprise MLOps strategy
- Fast data scoring performance streaming, real-time, and batch applications
- Model Deployment, Management, and Monitoring Services
- Built-in cognitive text services (English, French, Spanish, Italian)
- Pay only for actual scoring compute

OML Services Deployment

Enable key elements of an overall enterprise MLOps strategy

- Deploy native in-database OML Models
- Deploy third-party (ONNX format) models
- Score/inference using classification, regression, clustering, and feature extraction models
  - Synchronous single and small batch scoring
  - Asynchronous (large) in-database batch scoring
- Use cognitive image functionality by deploying ONNX-format third-party models with the ability to score using images or tensors

OML Services Model Management

Enable key elements of an overall enterprise MLOps strategy

- Store and manage in-database and ONNX-format models along with their metadata using REST endpoints
- Retrieve models for loading back to the database for use from SQL/R/Python APIs
- Create scoring endpoints for registered models with singleton, small batch, and large batch scoring
- Supports classification, regression, clustering, and feature extraction models
- Store, version, compare ML models
- Organize models within namespaces
- Import ONNX-format models produced from popular third-party packages like Tensorflow, PyTorch, MXNet, Scikit-learn, etc.

OML Services Monitoring

Enable key elements of an overall enterprise MLOps strategy

Data Monitoring

- Data drift detection
- Track and report on significant changes in the statistical properties of data
- Maintain data quality standards
- Notify users when data drift has occurred to take possible action

Model Monitoring

- Model concept and quality drift detection
- Track and report on, e.g., changes in prediction distribution and model accuracy, which may signal the need to rebuild a model or investigate causes
- Support in-database models
- Notify users when model drift has occurred to take possible action

OML Services Batch Scoring - *New!*

What is batch scoring and why is it important?

- In-database, scoring of multiple data records asynchronously, e.g., a group of or all customers, employees, or equipment
- Data is typically accessed through a database table or view
- Used in deployment scenarios where, for example, customer behavior changes over days, weeks, months, and years
- Score data regularly as its profile and behavior changes over time
- Use cases include lead generation, churn, demand forecasting, supply chain optimization, and many others
- Support for regression, classification, clustering, and feature extraction

Build models with PL/SQL and score with SQL

Build a classification model to predict customer acceptance to affinity card loyalty program

```
DECLARE
  v_setlst DBMS_DATA_MINING.SETTING_LIST;
BEGIN
  v_setlst('PREP_AUTO') := 'ON';
  v_setlst('ALGO_NAME') := 'ALGO_GENERALIZED_LINEAR_MODEL';
  DBMS_DATA_MINING.CREATE_MODEL2(
        'GLM_CLASS_MODEL',
        'CLASSIFICATION',
        'SELECT * FROM TRAIN_DATA',
         v_setlst,
        'CUST_ID',
        'AFFINITY_CARD');
END;
```

Score and return customers > 50% likely to be affinity card responders

```
SELECT * FROM (
 SELECT CUST_ID,
    PREDICTION_PROBABILITY(GLM_CLASS_MODEL,
      '1',
      USING A.*) PROBABILITY
  FROM TEST_DATA A)
WHERE PROBABILITY > 0.5;
```

MODEL NAME identifies the model in OML and makes the model available for SQL scoring

Oracle Machine Learning Services - Methods

Components with built-in Oracle Machine Learning

Admin

POST

- Token using ADB user and password

Generic

GET

- Metadata for all Versions: Version 1 Metadata
- Open API Specification

Repository

POST

- Store Model
- Update Model Namespace

GET

- Model Listing
- Model Info
- Model Metadata
- Model Content

DELETE

- Model

Deployment

POST

- Create Model Endpoint
- Score Model using Endpoint

GET

- Endpoints
- Endpoint Details
- Open API Specification for Endpoint

DELETE

- Endpoint

Asynchronous Jobs

POST

- Submit Job
- Update Job
- Perform Job Actions

GET

- Jobs Listing
- Job Details

DELETE

- Delete Job

Cognitive Text

POST

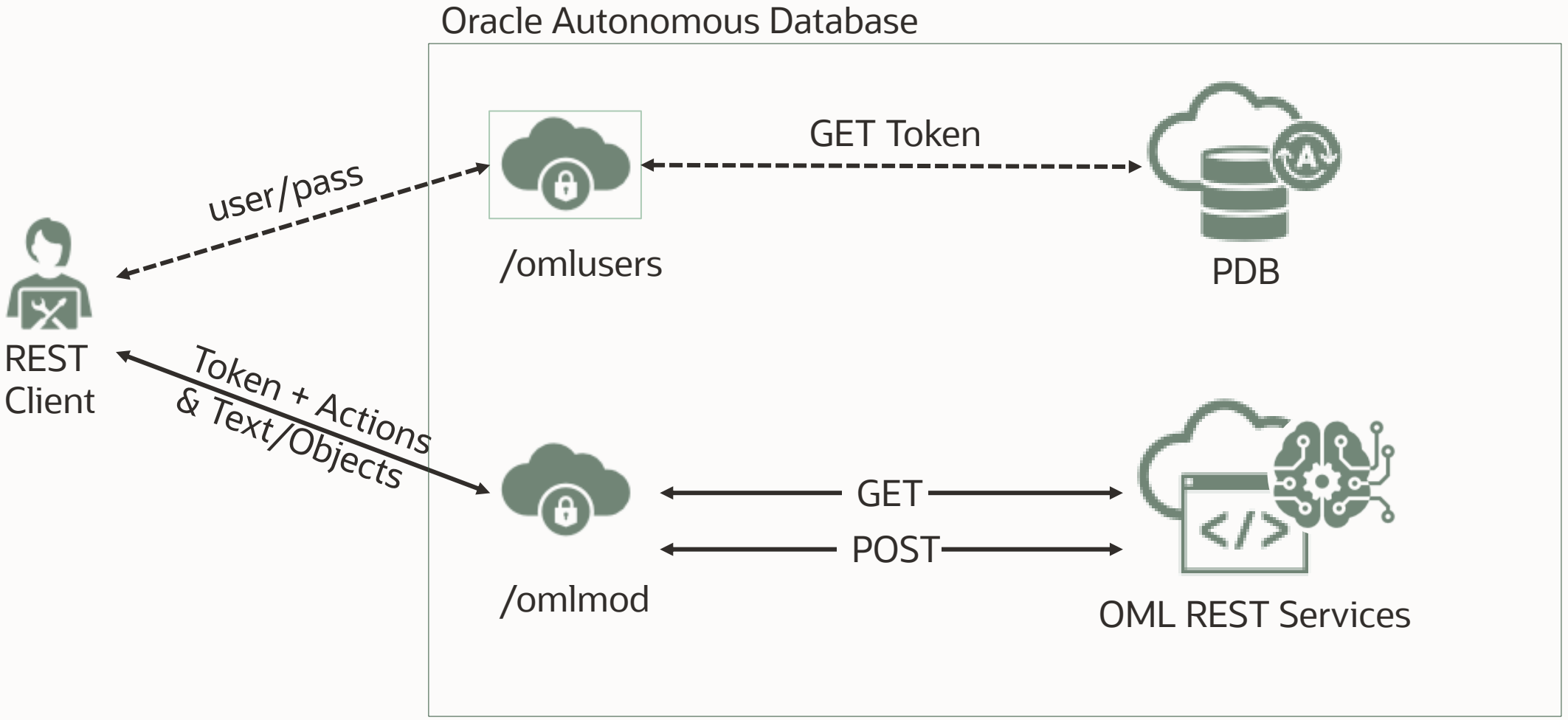
- Get Most Relevant Topics
- Get Most Relevant Keywords
- Get Summaries
- Get Sentiments
- Get Semantic Similarities
- Numeric Features

GET

- Get Endpoints

OML Services REST Architecture

Connectivity and use from client



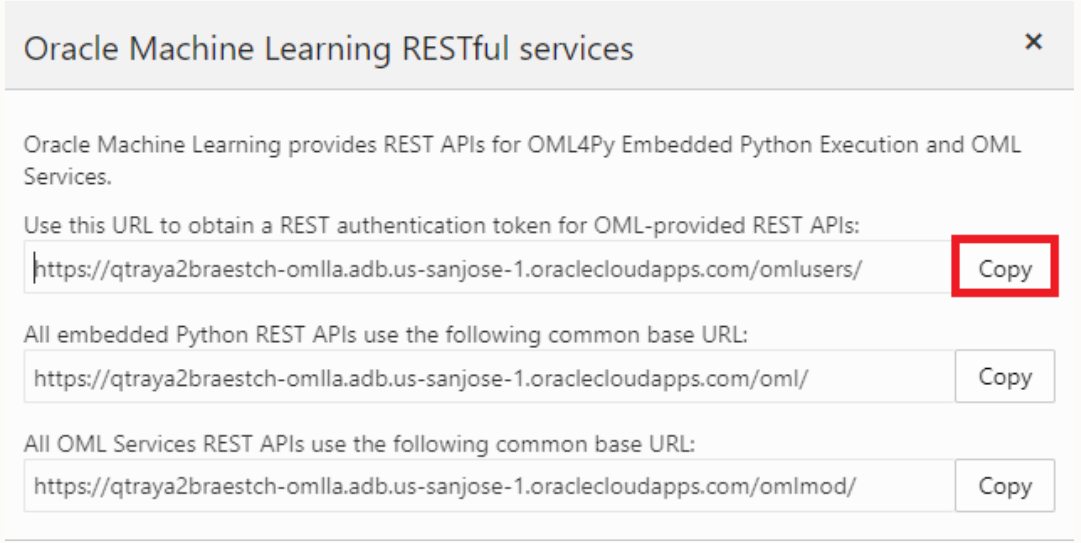
Oracle Machine Learning RESTful URLs

Where can I find the URLs that correspond to my tenancy?

Location of REST URLs

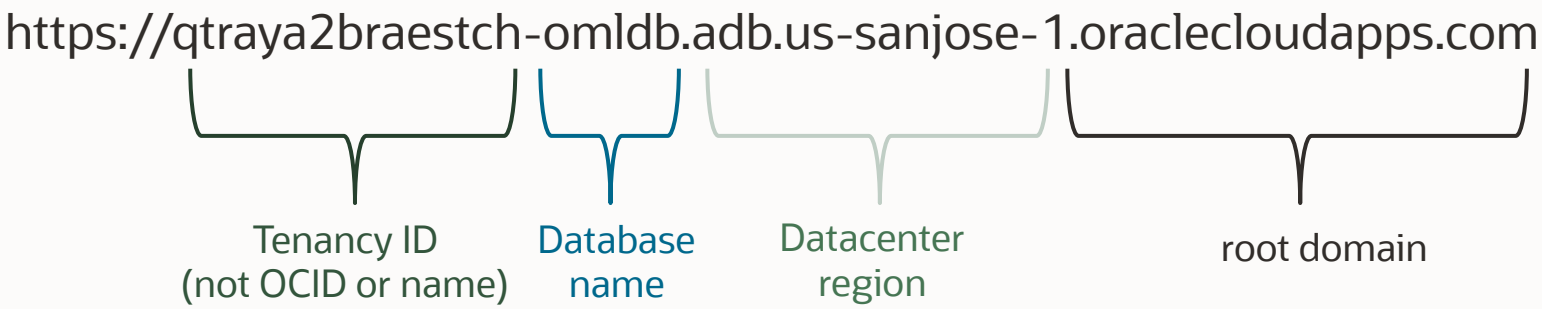
From your Oracle Autonomous Database instance:

1. Click **Database Actions**
2. Scroll down to **Oracle Machine Learning RESTful Services** under **Related Services** and copy the URL



OML Service REST URL

Standard URL for all OML REST endpoints



Same for Token acquisition

POST

Request a token

Initial call to get a token to access all other OML REST endpoints

To request a token for accessing OML endpoints, you need a valid user and password for your Oracle Autonomous Database user

Ensure OML administrator grants OML Developer privileges

For the following REST call, consider (see the previous slide for example):

`OML Service URL = tenancy/database/region/oraclecloudapps.com`

```
$ curl -X POST \
  --header 'Content-Type: application/json' \
  --header 'Accept: application/json' \
  -d '{"grant_type":"password", "username": "YourOMLuser", "password": "YourOMLpass"}'\
  "OML service URL/omlusers/api/oauth2/v1/token"
```


POST

Batch Scoring

Submit a batch scoring job – job schedule

```
$ curl -X POST "${omlservice}/omlmod/v1/jobs"
       -H "Authorization: Bearer ${token}"
       --data '{
        "jobSchedule":{
        "jobStartDate":"2023-04-10T10:15:00Z",
        "jobEndDate":"2023-04-15T10:15:00Z",
        "repeatInterval":"FREQ=DAILY",
        "maxRuns": "15"
         }...
```

Token variable

Job start and end dates

Job frequency

Maximum number of runs within the job time period

POST

Batch Scoring

Submit a batch scoring job – job properties

```
"jobProperties": {
  "jobname": "MY_BATCH_SCORING_JOB",
  "jobType": "BATCH_SCORING",
  "modelId": "35a97c7b-0ff4-4940-96f5-bfb29f64d223",
  "inputData": "CUSTOMERS360",
  "outputData": "CLASS_PRED1",
  "supplementalColumnNames":["CUST_ID","AFFINITY_CARD"],
  "jobServiceLevel": "MEDIUM",
  "topN": 2,
  "topNDetails": 2,
  "recompute": "true"}
```

- Job name and type
- Model ID
- Input/output data
- Columns from input data to identify output rows
- topN probabilities and prediction details
- Replace output table

GET

List Available Jobs

Call to get the available jobs

A list of jobs previously created by the OML user associated with the token is returned.

For the following REST call, consider:

`OML Service URL = tenancy/database/region/oraclecloudapps.com`

Remember to provide the full Token after "Bearer"

```
$ curl –X GET ‘OML Service URL/omlmod/v1/jobs’ \
 --header 'Content-Type: application/json’ \
 --header 'Accept: application/json' \
 --header 'Authorization: Bearer eyJhbGciOiJSUzI1NiJ9.....=='
```

This is the token

GET

Batch scoring job details

Call to get the batch scoring job details

Details for the job previously created by the OML user associated with the token is returned.

For the following REST call, consider:

```
OML Service URL = tenancy/database/region/oraclecloudapps.com
jobid = the Job ID for the batch scoring job
```

Remember to provide the full Token after "Bearer"

```
$ curl -X GET 'OML Service URL/omlmod/v1/jobs/jobid' \
 --header 'Content-Type: application/json' \
 --header 'Accept: application/json' \
 --header 'Authorization: Bearer eyJhbGciOiJSUzI1NiJ9.....=='
```

This is the token

Demo

Batch Scoring with OML Services using cURL and Postman

Poll #2: Understanding

How has this session helped your understanding of using OML Services for batch scoring?

- I learned what I need and plan to use OML Services for batch scoring
- I learned a lot and will look for opportunities to apply this in my own projects or with customers
- It was good to learn how OML Services supports a growing number of features, but I don’t have immediate needs

For more information…

OML Webpage

https://oracle.com/machine-learning

OML Services Batch Scoring Blog

https://bit.ly/3muLox1

OML GitHub Repository

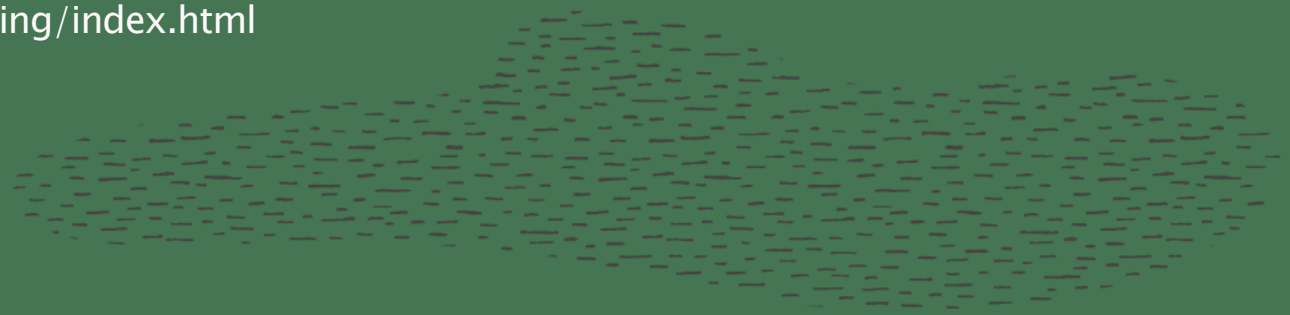
https://bit.ly/omlgithub

OML Office Hours

https://bit.ly/omlofficehours

OML Documentation

https://docs.oracle.com/en/database/oracle/machine-learning/index.html



Q & A

Thank you

sherry.lamonica@oracle.com

miles.novotny@oracle.com