

ORACLE



**This session will
be recorded**

Oracle Machine Learning

AskTOM Office Hours – Feature Highlight

Feature Highlight: Data and Model Monitoring in Autonomous Database
with Miles Novotny and Sherry LaMonica

Host: Mark Hornick

Product Management, Oracle Machine Learning

Feature Highlight: Data and Model Monitoring in Autonomous Database

Speakers: Miles Novotny and Sherry LaMonica

Machine Learning models are increasingly used to make business decisions in production environments. It is critical to ensure that these models make accurate predictions and are robust to shifts in the data.

Oracle Machine Learning Services extends OML functionality in Autonomous Database to support model deployment and model lifecycle management via REST APIs. In the coming weeks, data and model monitoring capabilities will be introduced. With these features, users can track the accuracy of their machine learning models in production and understand data trends by viewing performance degradation alerts. Join us to learn about these upcoming features and see demonstrations using cURL and Postman.

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, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.

Poll #1: Usage

What is your level of familiarity with data monitoring and model monitoring?

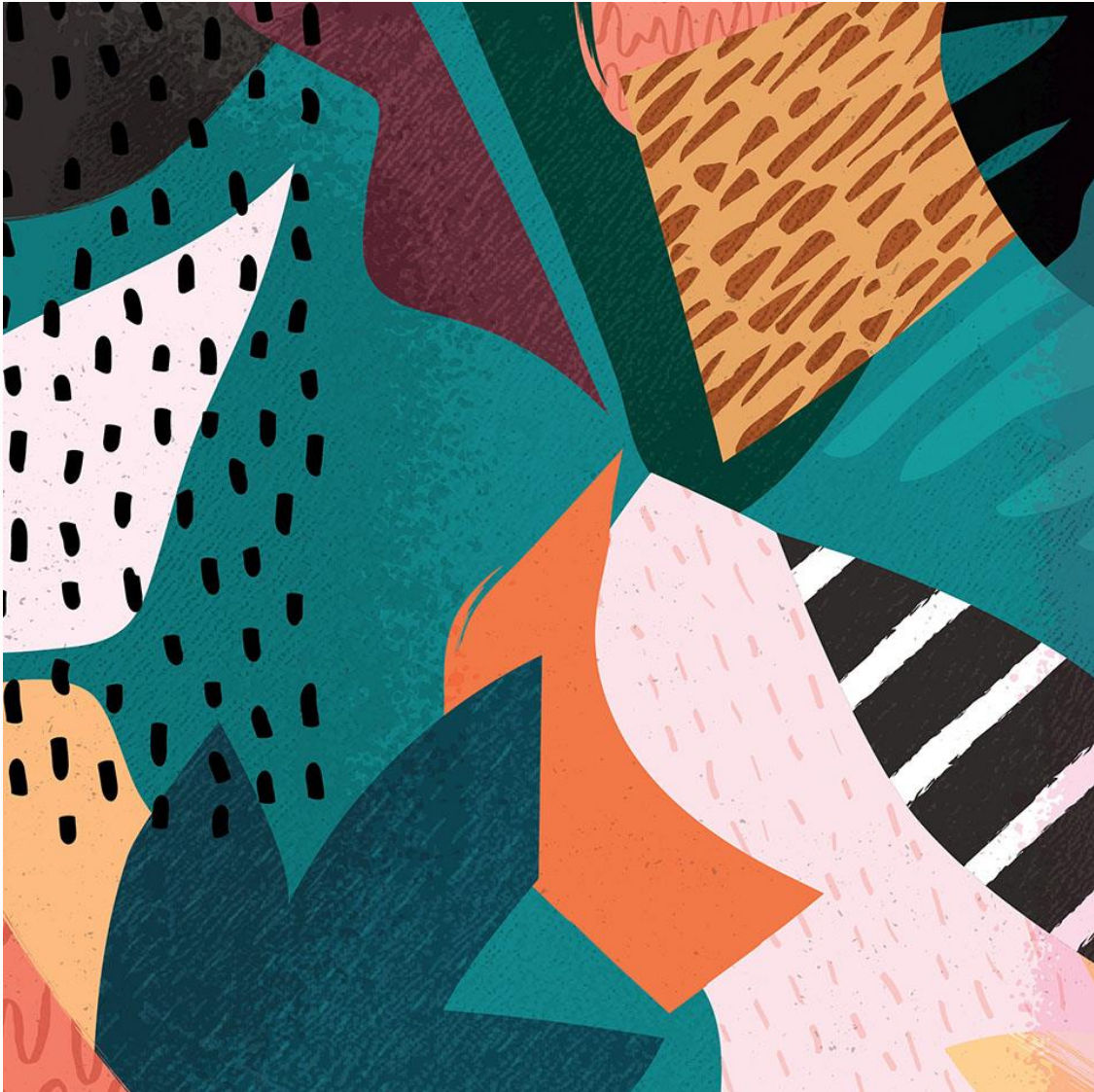
- I have monitored data and models
- I have monitored data
- I have monitored models
- I am familiar with the concepts, but have never used them
- I am interested to find out what these concepts are all about

ORACLE

Oracle Machine Learning

AskTOM Office Hours – Feature Highlight

Feature Highlight: Data and Model Monitoring in Autonomous Database
with Miles Novotny and Sherry LaMonica



1

What are data and model monitoring and why are they important?

2

Data and model monitoring with OML Services overview

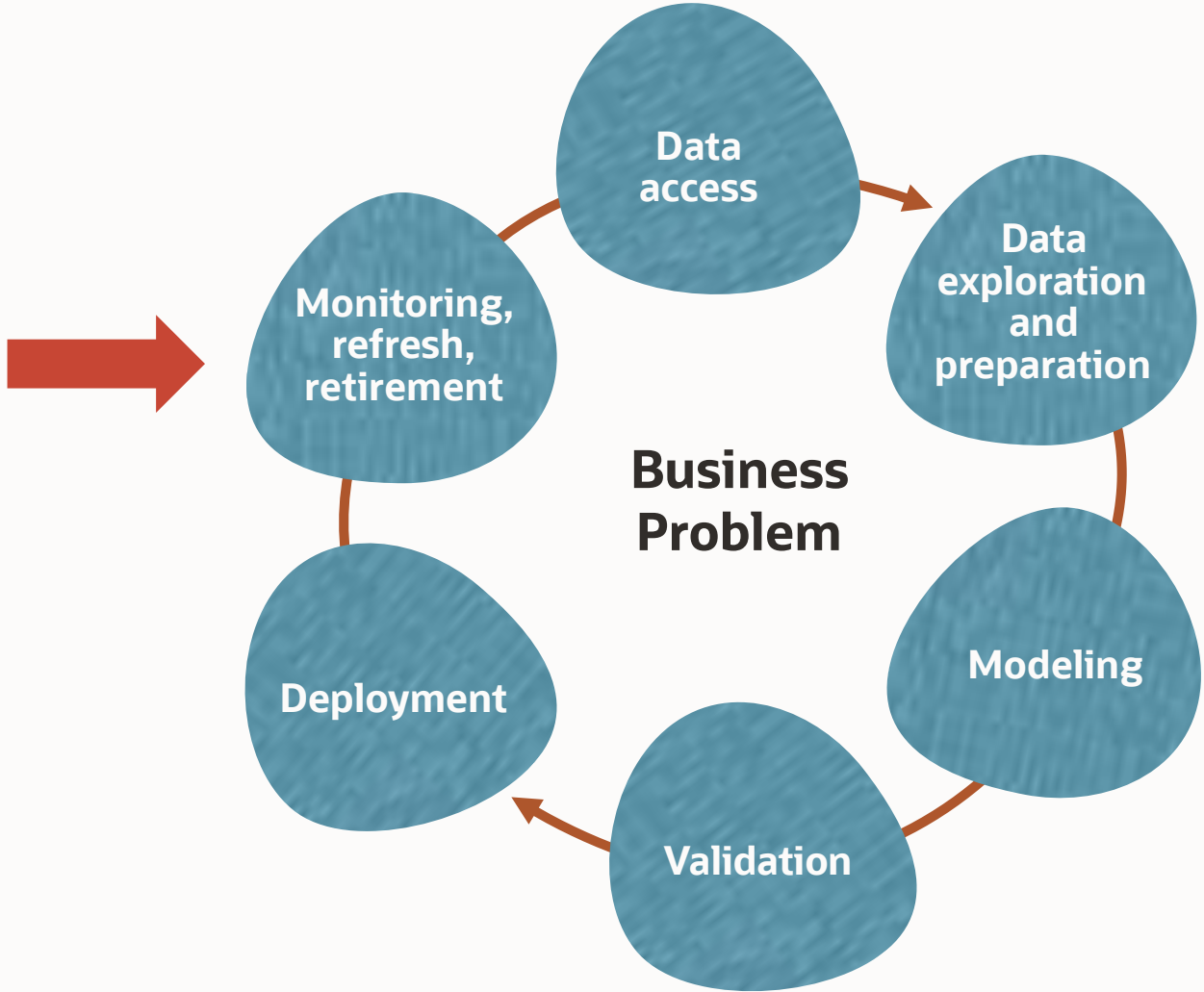
3

Data and model monitoring with OML Services demo

4

Q&A

Where Data Monitoring and Model Monitoring Happen



What is Data Drift?

Gradual change in data over time

What causes data drift?

- Changing business environment (e.g. legal or government changes)
- Evolving behavior (changes in customer preferences, regional trends)
- Data quality issues (incorrect or missing data)
- Issues in upstream data processing (sensor wear and tear)

What is Data Monitoring?

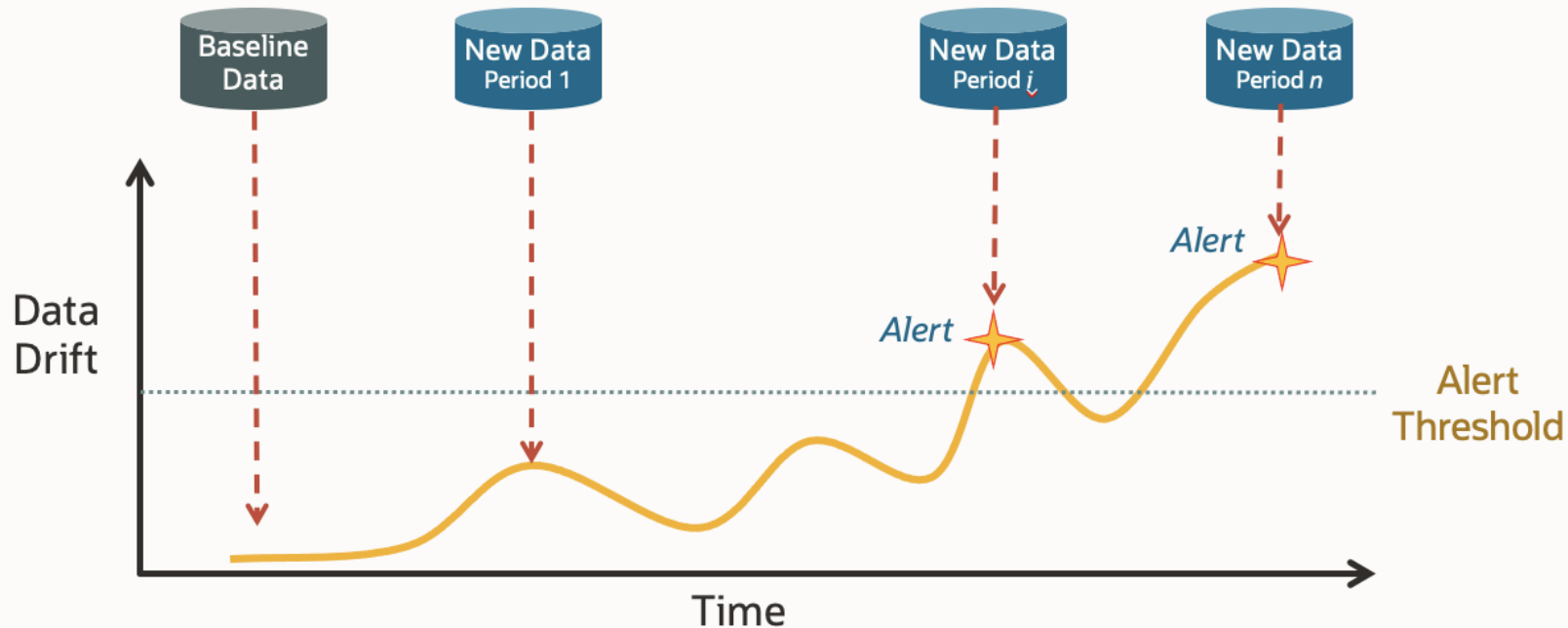
Track data over time to monitor potential data drift

How to define drift?

- Baseline data
- New data
- “Can we build a model to predict if a data row belongs to baseline or new data”?

How to measure drift?

- Look a changes in interactions between variables
- Bivariate interactions are not easily explored with descriptive statistics
- Focus on interactions of interest
- 0-1 scale based on AUC of model predicting between baseline and new data

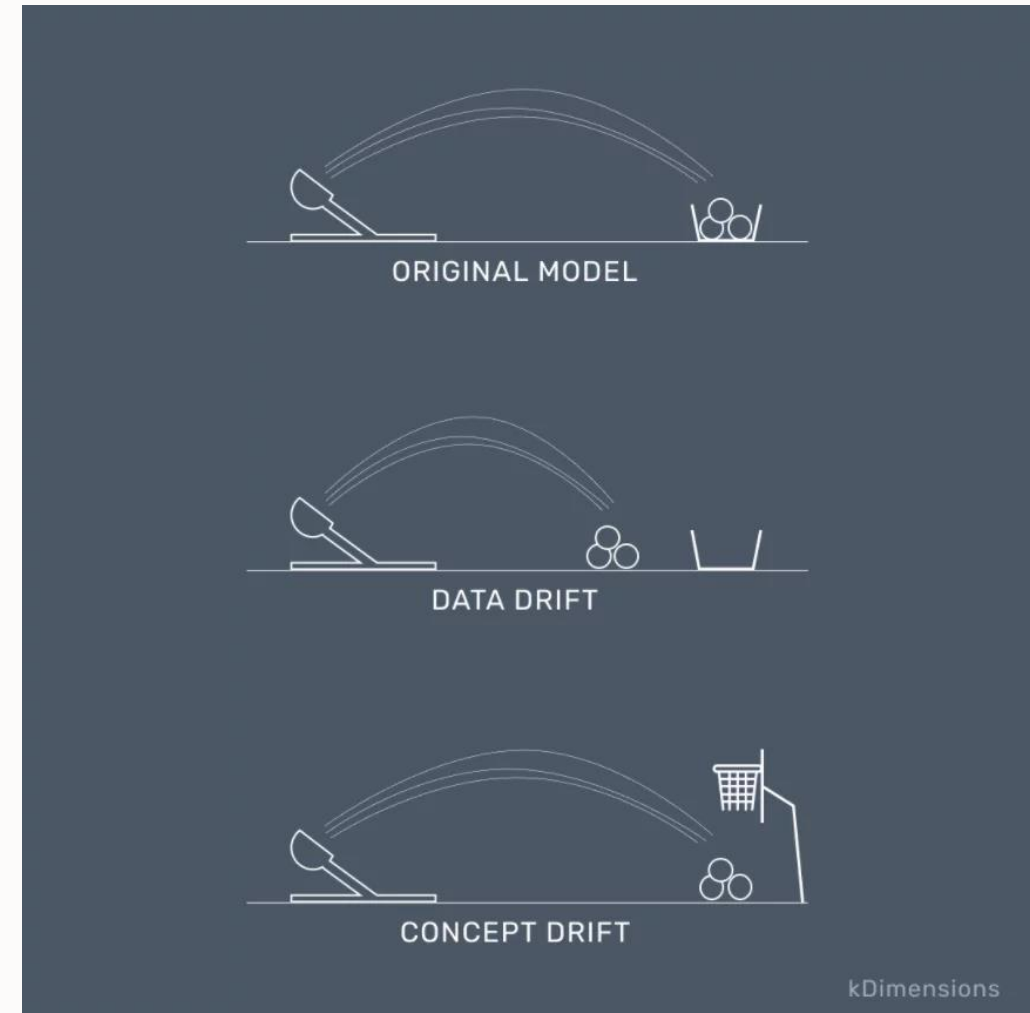


What is Model Drift?

Gradual degradation in prediction quality of a model

Caused by:

- Data drift: The input to our model has changed
- Concept drift: What we are trying to predict has changed

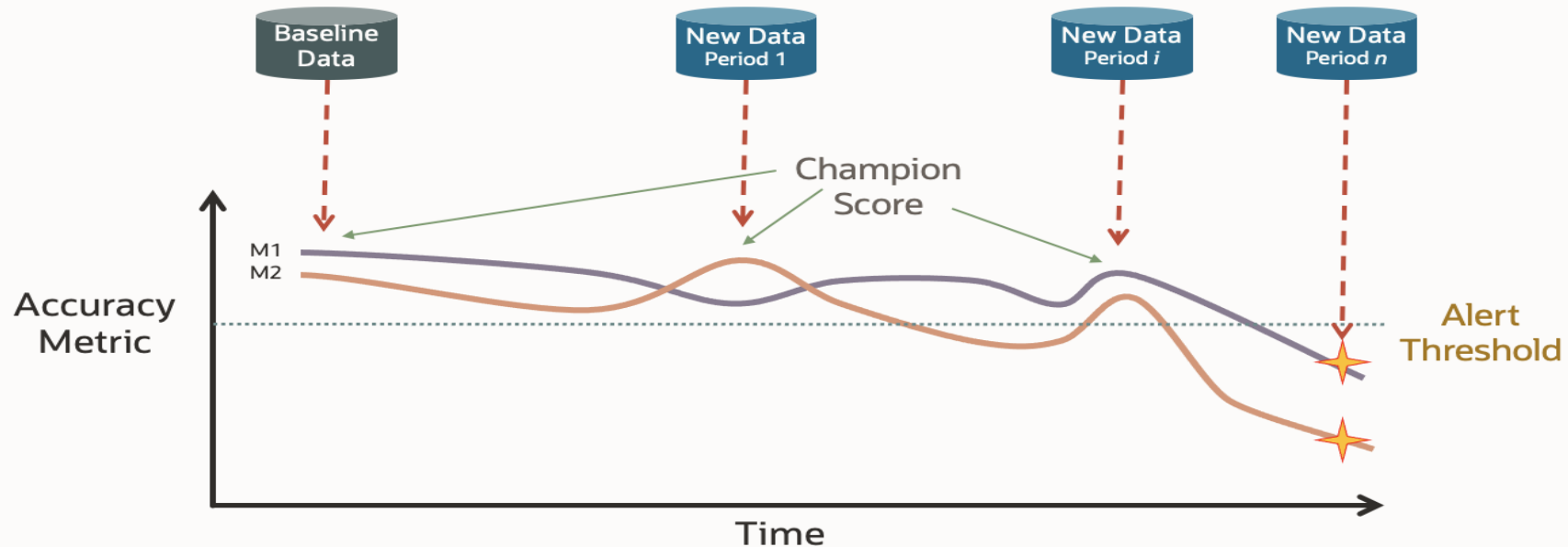


What is Model Monitoring?

Tracking prediction quality over time for potential performance degradation

- Baseline data
- New data
- Performance metrics (Accuracy, F1, MSE, MAE)
- 0-1 scale based on difference between baseline and new data metrics (scaled by baseline metric)

Data monitoring is a prerequisite to effective model monitoring



Data Monitoring/Model Monitoring Use Case

A bank uses ML to preapprove people for loans

Model monitoring tells them that the the quality of this model is degrading

Data monitoring results tell them that customer's data has not drifted significantly

Concept drift could be the cause

Looking at recent business environment changes, and note that the economy has improved as of late

This change has made them want to give out more loans, but the unchanged model was still approving at the previous rate

They retrain the model, taking these concept changes into account



Why is Data Monitoring Important?

Data monitoring as a prerequisite to effective model monitoring

Data monitoring for non-machine learning use cases:

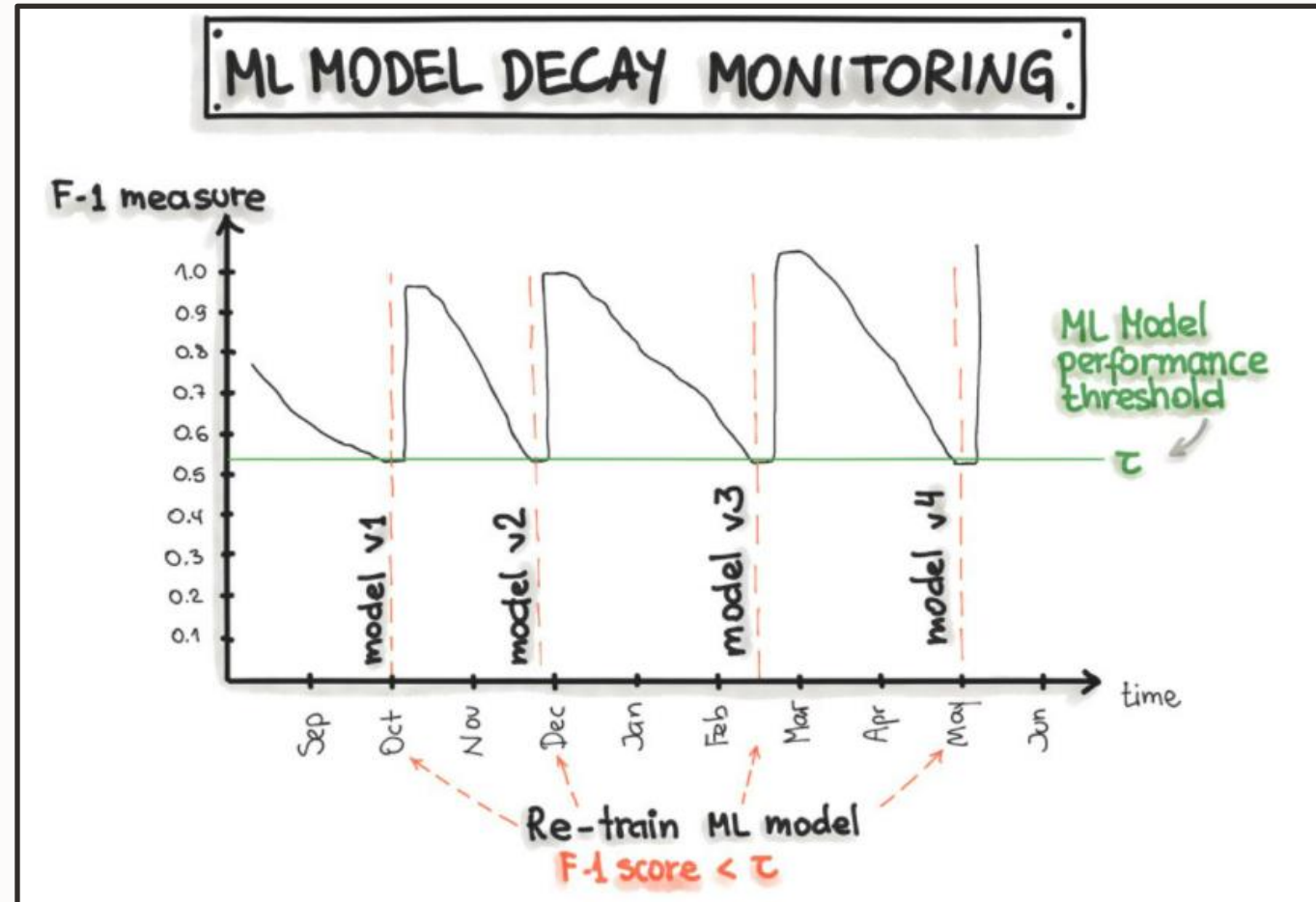
- Maintain data quality standards
- Analyze data when it passes a threshold
- Perform some action when data passes a threshold

Why is Model Monitoring Important?

Understand quality of deployed models on an hourly, daily, weekly basis

Know why models degrade (data, concept drift)

Gain awareness when production models degrade to maximize business value



Oracle Machine Learning Services overview

Extends OML functionality on Autonomous Database to support model deployment and model lifecycle management

The REST API for OML Services provides:

- Endpoints that enable the **storing machine learning models** along with its metadata
- Creation of **scoring endpoints** for the registered models.
- Support for **classification and regression of third-party models in ONNX format**, including from packages like Scikit-learn and TensorFlow, among others.
- **Proprietary cognitive text** capabilities in English, Spanish, French and Italian for topic discovery, keywords, summary, sentiment, and feature extraction, based on a Wikipedia knowledge base using Embeddings.
- **Cognitive image functionality**, supported through the ONNX format third-party model deployment feature, with the ability to score using images or tensors.
- **COMING SOON!** Asynchronous jobs for **batch scoring** and **model management with data and model monitoring capabilities** to assess the accuracy of models deployed in production and detect potential degradation

Oracle Machine Learning Services - Methods

Components with built-in Oracle Machine Learning



| Admin | Repository | Deployment | Asynchronous Jobs | Cognitive Text |
|---|--|---|---|---|
| POST <ul style="list-style-type: none">Token using ADB user and password | POST <ul style="list-style-type: none">Store ModelUpdate Model Namespace | POST <ul style="list-style-type: none">Create Model EndpointScore Model using Endpoint | POST <ul style="list-style-type: none">Submit JobUpdate JobPerform Job Actions | POST <ul style="list-style-type: none">Get Most Relevant TopicsGet Most Relevant KeywordsGet SummariesGet SentimentsGet Semantic SimilaritiesNumeric Features |
| Generic | | | | |
| GET <ul style="list-style-type: none">Metadata for all Versions: Version 1 MetadataOpen API Specification | GET <ul style="list-style-type: none">Model ListingModel InfoModel MetadataModel Content | GET <ul style="list-style-type: none">EndpointsEndpoint DetailsOpen API Specification for Endpoint | GET <ul style="list-style-type: none">Jobs ListingJob Details | |
| | DELETE <ul style="list-style-type: none">Model | DELETE <ul style="list-style-type: none">Endpoint | DELETE <ul style="list-style-type: none">Delete Job | GET <ul style="list-style-type: none">Get Endpoints |



Model Monitoring Performance Metrics

Metrics available to measure model drift

Regression

Mean Squared Error

Mean Absolute Error

Mean Squared Log Error

Median Absolute Error

Classification

Accuracy

Balanced accuracy

AUC

F1

Precision

Recall

Data Monitoring Performance Metrics

Metrics available to measure data drift

Numeric Variables

Mean

Standard deviation

Range (minimum, maximum)

Number of nulls



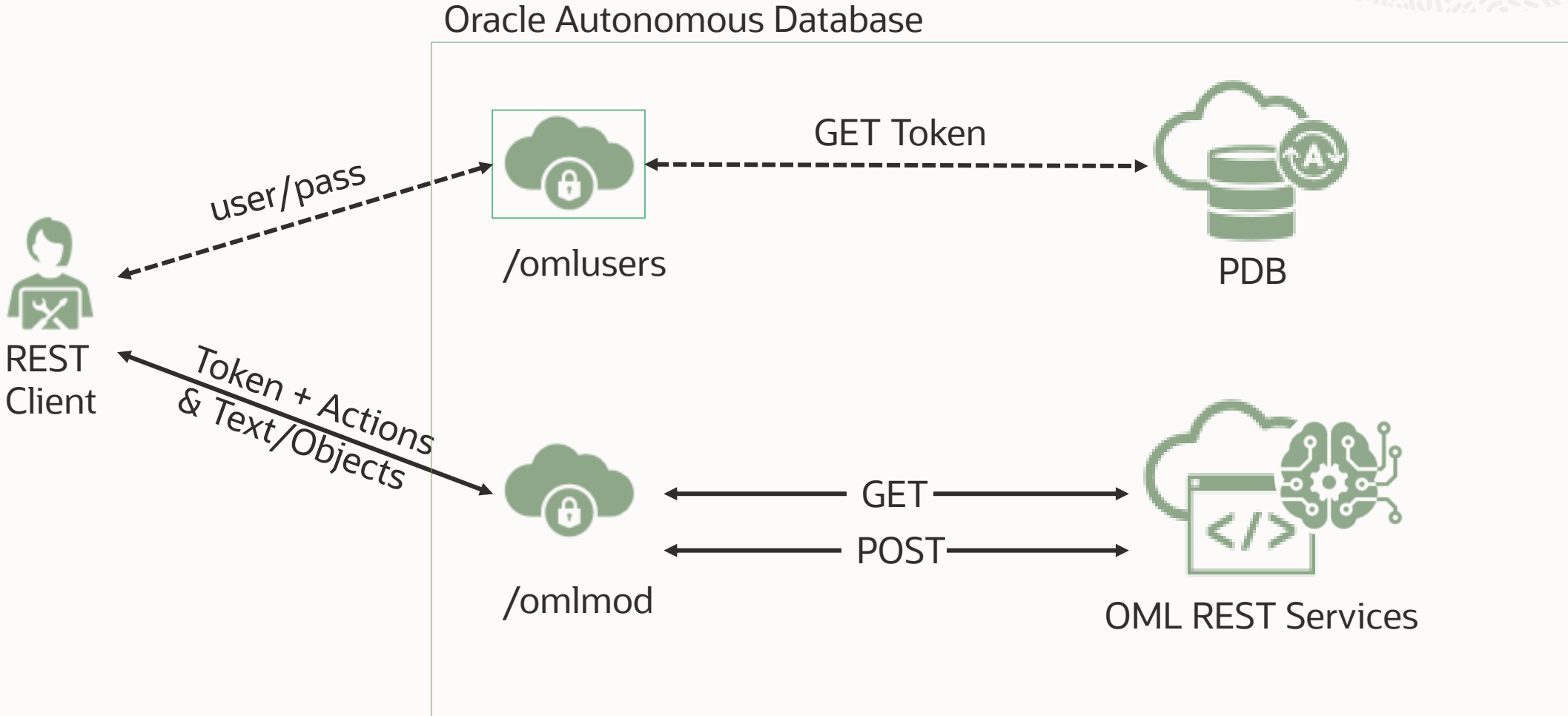
Categorical Variables

Number of unique values

Number of nulls

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

Oracle Machine Learning User Interface

Oracle Machine Learning provides several components accessible through a common user interface. OML Notebooks supports Python, SQL, PL/SQL, and Markdown interpreters, with access to in-database ML through OML4Py and OML4SQL. OML Models supports managing and deploying in-database models. OML AutoML UI provides a no-code user interface to build, evaluate, and deploy in-database models using automated machine learning.

RESTful Services and SODA

Oracle REST Data Services (ORDS) provides HTTPS interfaces for working with the contents of your Oracle Database in one or more REST enabled schemas.

All ORDS delivered resources for your Autonomous Database will have the following common base URL:

<https://qtraya2braestch-omldb.adb.us-sanjose-1.oraclecloudapps.com/ords/>

Copy

Oracle Machine Learning RESTful services

Oracle Machine Learning provides REST APIs for OML4Py Embedded Python Execution and OML Services.

Use this URL to obtain a REST authentication token for OML-provided REST APIs:

<https://qtraya2braestch-omldb.adb.us-sanjose-1.oraclecloudapps.com/omlusers/>

Copy

All embedded Python REST APIs use the following common base URL:

<https://qtraya2braestch-omldb.adb.us-sanjose-1.oraclecloudapps.com/oml/>

Copy

All OML Services REST APIs use the following common base URL:

<https://qtraya2braestch-omldb.adb.us-sanjose-1.oraclecloudapps.com/omlmod/>

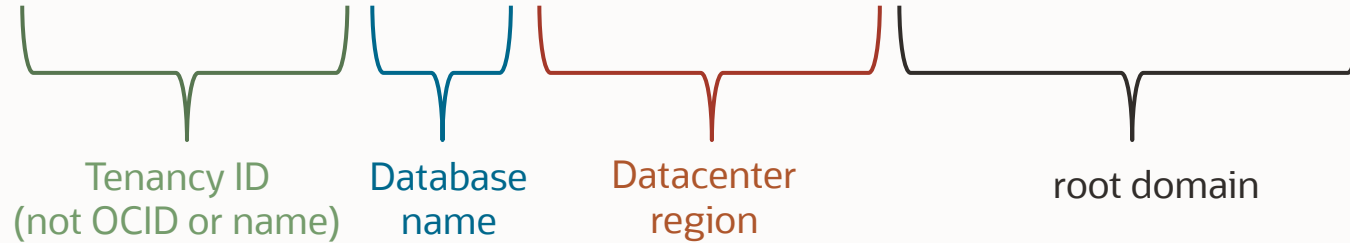
Copy

OML Service REST URL

Standard URL for all OML REST endpoints



https://qtraya2braestch-omldb.adb.us-sanjose-1.oraclecloudapps.com



Same for Token acquisition



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"
```



Data Monitoring

Submit a data monitoring job – job schedule

```
$ curl -X POST "${omlservice}/omlmod/v1/jobs"  
--H "Authorization: Bearer ${token}"  
--data '{  
  "jobSchedule": {  
    "jobStartDate": "2023-03-01T10:15:00Z",  
    "jobEndDate": "2023-03-15T10:15:00Z",  
    "repeatInterval": "FREQ=DAILY",  
    "maxRuns": "15"  
  },  
  "jobProperties": {continued..
```

Token variable

Job start and end dates

Job frequency

Maximum number of runs

Data Monitoring

Submit a data monitoring job – job properties

```
"jobProperties": {  
  "jobname": "MY_DATA_MONITOR_JOB",  
  "jobType": "DATA_MONITORING",  
  "outputData": "DATA_MONITOR_OUT",  
  "baselineData": "IRIS_DM_BASE",  
  "newData": "IRIS_DM_NEW",  
  "threshold": 0.75,  
  "timeColumn": "TIMECOL",  
  "startDate": "2020-01-15T19:30:05Z",  
  "endDate": "2021-01-15T19:30:05Z",  
  "featureList": ["SEPAL.LENGTH", "SEPAL.WIDTH"],  
  ... }  
}
```

Job name and type

Output table

Drift threshold

Monitoring start and end date
and time using TIMECOL in
newData

Features to be monitored

Model Monitoring

Submit a model monitoring job – job schedule

```
$ curl -X POST "${omlservice}/omlmod/v1/jobs"  
--H "Authorization: Bearer ${token}"  
--data '{  
  "jobSchedule": {  
    "jobStartDate": "2023-03-01T10:15:00Z",  
    "jobEndDate": "2023-03-15T10:15:00Z",  
    "repeatInterval": "FREQ=DAILY",  
    "maxRuns": "15"  
  },  
  "jobProperties": {continued..
```

Token variable

Job start and end dates

Job frequency

Maximum number of runs

Model Monitoring

Submit a model monitoring job – job properties



```
"jobProperties": {  
  "jobname": "MY_MODEL_MONITOR_JOB",  
  "jobType": "MODEL_MONITORING",  
  "outputData": "MODEL_MONITOR_OUT",  
  "baselineData": "IRIS_MM_BASE",  
  "newData": "IRIS_MM_NEW",  
  "threshold": 0.75,  
  "timeColumn": "TIMECOL",  
  "startDate": "2020-01-15T19:30:05Z",  
  "endDate": "2021-01-15T19:30:05Z",  
  "performanceMetric": "BALANCED_ACCURACY",  
  "modelList": ["b3060ec0-45a3-4819-825f-14082afd0f44"],  
  ... }  
}
```

Job name and type

Output table

Drift threshold

Monitoring start and end date
and time using TIMECOL in
newData

Performance metric

Model or list of model IDs

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

Data monitoring job details

Call to get the data monitoring 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 model monitoring 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

Model monitoring job details

Call to get the model monitoring 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 model monitoring 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

Poll #2: Future use

What is your likelihood of using data or model monitoring in your future projects?

- I plan to use data and model monitoring in future projects
- I plan to use data monitoring in future projects
- I plan to use model monitoring in future projects
- I need more information before I can start using these features
- I don't plan to use these functionalities

Q & A

Thank you



sherry.lamonica@oracle.com



miles.novotny@oracle.com

