

# Data and Model Monitoring – The Step <u>Not</u> To Skip In Solution Deployment

Mark Hornick Senior Director, Oracle Al and Machine Learning Product Management



Copyright © 2025, AnDOUC. All rights reserved.

# Future & Past TechCasts:



Data and Model Monitoring – The Step Not To Skip In Solution Deployment

Presented by Mark Hornick

May 29th

Thwart Toil Through Tiles: Leveraging Oracle 23ai's Latest Geospatial Features

Presented by Jim Czuprynski



Latest Features of OAC: April 2025 Release

Presented by Dan Vlamis, Wayne van Sluys, & Tim Vlamis

## **TechCast Archive**

2025		2024	2023	2022	2021		2020
Date	Title			Presenter(s)		Replay	Download(s)
Apr 3	Exploring Relationships in Your Data With Oracle Analytic Cloud (OAC)			Melli Annamalai, Philippe Lions & Gautam Pisharam		Video	Slides
Mar 6	Automating Oracle Analytics Cloud Administration with REST APIs			Joel Acha		Video	Slides
Feb 6	Our Favorite New Features in OAC: February 2024 and January 2025 Releases			Dan Vlamis, Wayne Van Sluys, Philippe Lions	Gautam Pisharam,	Video	Slides
Jan 23	Leveraging Vector Search for RAG in Generative AI		Kai Yu		Video	Slides	
Jan 9	The Oracle AI	Microservices Sandbox for RA	G Rapid Prototyping	Corrado De Bari, Mark Nelson,	& John Lathouwers	Video	Slides

Submit a topic to share at https://andouc.org/techcasts/



www.andouc.org





## Let's Connect





#### ORACLE



# Data and Model Monitoring

## The step not to skip in solution deployment

### Mark Hornick Senior Director, Oracle Al and Machine Learning Product Management

May 1, 2025

### Agenda

- ML process
- Why monitor?
- Data monitoring
- Model monitoring
- Data bias detection

# ML Process

Monitoring... the step not to skip



## Machine learning process - novice approach

Linear approach with solution deployment end-goal



0

## Machine learning process – experienced approach



Cross-industry standard process for data mining (CRISP-DM)

# Why monitor?

### The importance of data monitoring

#### **Diet Pills -2016**

#### **Clinical trials**

Without careful monitoring, harmful side effects or lack of efficacy may go unnoticed – putting participants at risk, leading to faulty conclusions.

#### Telecom churn – 2008-13

### **Customer service**

Unmonitored customer satisfaction metrics can lead to losing customers to competitors without understanding why.

#### Manufacturing

Not monitoring product quality at each step can result in defects making it to customers, harming the brand reputation and needing costly recalls

#### Supply chain

Ignoring delivery times, costs, and supplier quality hides inefficiencies and risks customer satisfaction not receiving orders on time

**Baby formula shortage – 2022** 

#### **Treadmills - 2021**

Diligent monitoring provides visibility and informs course correction

### 0

## The importance of ML model monitoring

ML models often need ongoing attention to maintain their value

#### Model performance degradation

As new data arrives, distributions and patterns may change, which warrants model retraining

#### New data may reveal biases

New data may have different biases than the original, resulting in model prediction unfairness

#### **Prediction errors point to data issues** Increase in certain prediction errors can indicate systemic data problems, indicating pipeline issues

**Regulation/compliance requirements** Regulated industries like banking must comply with fairness, transparency, and accountability rules

# Data Monitoring

**Data drift** What type of changes are important?

Change in number of distinct values: 3 categories  $\rightarrow$  5 categories

Numeric feature distribution drift







## Data Monitoring UI

No-code user interface for tracking changes in data over time

Capture changes not obvious from individual attribute statistics

Explore changes in feature interactions efficiently

Benefit from ML to model complex interactions and discover nonobvious data relationships

Examine statistics of most impactful features and their interactions for insight into data changes



### **OML** Data Monitoring no-code user interface



### **OML** Data Monitoring no-code user interface





# The Dataset

Individual Household Electric Power Consumption

Temporal Measurements Date: calendar day measurement taken Time: time when measurement was taken Specific Area Electricity Measurements Sub-metering 1: Watt-hours used in kitchen Sub-metering 2: Watt-hours used in the laundry room Sub-metering 3: Watt-hours used by water heater and A/C

#### **Main Power Measurements**

Global Active Power: actual kilowatts consumed at the time of the reading
 Global Reactive Power: electricity that doesn't perform direct work but is needed for some appliances (e.g., those with motors or compressors) - measured in kilowatts
 Voltage: how strongly electricity is being pushed through the home's wiring system
 Global Intensity: total flow of electricity (amps) through the home at the time of the reading

### Using the data



0

# Demonstration

**OML Data Monitoring UI** 

## Data Monitoring using REST API

cURL examples using OML Services



Get Access Token	<pre>\$ curl -X POST header 'Content-Type: application/json' header 'Accept: application/json' data '{"grant_type":"password", "username":"<username>", "password":"<password>"}' "<oml service="" url="">/omlusers/api/oauth2/v1/token"</oml></password></username></pre>
Create data monitoring job	<pre>\$ curl -X POST "<omlserviceurl>/omlmod/v1/jobs" header "Authorization: Bearer <token>" header 'Content-Type: application/json' data '{"jobSchedule": {"jobStartDate": "2025-03-24T20:30:26Z",}, "jobProperties": {"jobName": "PowerDataMonitor",}, "jobType": "DATA_MONITORING", }' # Response contains job ID: OML\$73BB5308_166A_4D81_99BACEA599D1</token></omlserviceurl></pre>
View job details using job ID	<pre>\$ curl -X GET "<oml service="" url="">/omlmod/v1/jobs/OML\$73BB5308_166A_4D81_99BACEA599D1" header 'Accept: application/json' header 'Content-Type: application/json' header "Authorization: Bearer <token>"</token></oml></pre>
View results using SQL	<pre>SELECT START_TIME, END_TIME, IS_BASELINE, THRESHOLD, HAS_DRIFT, round(DRIFT, 3),     FEATURE_NAME, ROUND(IMPORTANCE_VALUE, 3) FROM OML\$73BB5308_166A_4D81_99BACEA599D1_PowerDataMonitor ORDER BY FEATURE_NAME, IS_BASELINE DESC</pre>

# Model Monitoring

### **Model drift** Contributing factors



## **Concept drift**

Changing patterns in data make ML model less accurate





## Model Monitoring UI

No-code user interface for tracking changes in model quality over time

Help detect deteriorating model accuracy and provide insights into underlying causes

Observe drift in distribution of predictions where known target values are not readily available to measure accuracy

Depict individual predictors in terms of model feature impact and drift models

Use model accuracy trends and model feature impact to guide model rebuilding



₩ 🖩

11/11/10 11/18/10 11/25/10



### OML Model Monitoring no-code user interface

ORACLE Machine Learning     Demos 2025     OMLUSER Workspace							OMLUSER 🔻	?
> What's New						> R	lecent Notebook:	s
V Quick Actions						> R	lecent Experimen	nts
AutoML Create and run AutoML Experiments	Models Manage and Deploy Machine Learning Models	Data Monitors Monitor Data Drift	<b>Model Monitors</b> Monitor Model Drift	<b>Scratchpad</b> Run Scratchpad	<b>Notebooks</b> The place for data discovery and analytics			
		<b>Jobs</b> Schedule notebooks to run at certain times	Examples Check out some examples					

### **OML Model Monitoring no-code user interface**



# Demonstration

**OML Model Monitoring UI** 

## Model Monitoring using REST API

### cURL examples using OML Services



Get Access Token	<pre>\$ curl -X POST header 'Content-Type: application/json' header 'Accept: application/json' data '{"grant_type":"password", "username":"<username>", "password":"<password>"}' "<oml service="" url="">/omlusers/api/oauth2/v1/token"</oml></password></username></pre>
Create data monitoring job	<pre>\$ curl -X POST "<omlserviceurl>/omlmod/v1/jobs" header "Authorization: Bearer <token>" header 'Content-Type: application/json' data '{"jobSchedule": {"jobStartDate": "2025-03-24T20:30:26Z",}, "jobProperties": {"jobName": "Power_Model_Monitor",}, "jobType": "MODEL_MONITORING', }' # Response contains job ID: OML\$736F509B_FC1A_4D0A_AC75_553F1D6C5D97</token></omlserviceurl></pre>
View job details using job ID	<pre>\$ curl -X GET "<oml service="" url="">/omlmod/v1/jobs/OML\$736F509B_FC1A_4D0A_AC75_553F1D6C5D97" header 'Accept: application/json' header 'Content-Type: application/json' header "Authorization: Bearer <token>"</token></oml></pre>
View results using SQL	SELECT IS_BASELINE, MODEL_ID, METRIC, HAS_DRIFT, DRIFT, MODEL_TYPE, THRESHOLD, MODEL_METRICS FROM OML\$736F509B_FC1A_4D0A_AC75_553F1D6C5D97_Power_Model_Monitor

# Data Bias Detection

## **Data Bias Detector**

OML Services REST API

Helps to support responsible AI initiatives

- Bias and fairness monitoring
- Help with accountability for AI systems

Identify possible bias in data early in the ML lifecycle, which can impact ML model quality

Assess if data bias may put certain groups at a disadvantage

Bias may arise due to, for example,

- Selection bias
- Survivorship bias
- Coverage bias

**Bias metrics** 

- Class Imbalance
- Statistical Parity
- Conditional Demographic Disparity

ne Learning Services Data Bias Detector





### **OML Services Data Bias Detector**

Metric property objectives: bounded, asymmetrical, distribution-free, computationally inexpensive

#### Class Imbalance (CI)

Evaluate mismatch between training data and population to which model will be applied

- Positive values  $\rightarrow$  group 1 has more training data
- Near zero  $\rightarrow$  groups are balanced in training data
- Negative values  $\rightarrow$  group 2 has more training data

### Statistical Parity (SP)

a.k.a. independence, group fairness, demographic parity, and disparate impact

- Positive values → outcomes where group 1 is accepted more than rejected
- Perfect zero score  $\rightarrow$  model does not predict any subgroup at a different rate than rest of population
- Negative values → outcomes where group 1 is rejected more than accepted

### Conditional Demographic Disparity (CDD)

Enables ruling out Simpson's paradox

- Positive values  $\rightarrow$  group 1 is accepted more than rejected
- Near zero  $\rightarrow$  no demographic disparity on average
- Negative values  $\rightarrow$  group 1 is rejected more than accepted

## **Conditional Demographic Disparity can identify Simpson's paradox**

Examples of Simpson's Paradox

Trend appears in several groups of data but disappears or reverses when the groups are combined

### UC Berkeley Gender Bias Case (1973)



Initially, it seemed male applicants had a higher acceptance rate than female applicants. However, when examining individual departments, no single department was significantly biased against women; in fact, women were favored in many departments. Overall bias due to women applying to more competitive departments with lower acceptance rates.

- Bickel, P. J., Hammel, E. A., & O'Connell, J. W. (1975)

### **Smoking and Mortality Rates**

In a study of smoking and mortality rates among British doctors, it was found that non-smokers had higher mortality rates than smokers when all ages were combined. However, when age groups were analyzed separately, smokers had higher mortality rates within each age group. The overall trend was because non-smokers tended to be older on average, and older people have higher mortality rates regardless of smoking status.

- Doll, R., & Hill, A. B. (1956)

## **Data Bias Detector using REST API**

### cURL examples using OML Services

Get Access Token	<pre>\$ curl -X POST \    header 'Content-Type: application/json'    header 'Accept: application/json'    data '{"grant_type":"password", "username":"<username>", "password":"<password>"}'     "<oml service="" url="">/omlusers/api/oauth2/v1/token"</oml></password></username></pre>
Create data bias	<pre>\$ curl -X POSTheader 'Content-Type: application/json'</pre>
detection job	
View job details using job ID	<pre>\$ curl -X GET "<oml service="" url="">/omlmod/v1/jobs/OML\$8017EE54_C305_47A5_9603_0EFA84B6B624" \    header 'Accept: application/json' \    header 'Content-Type: application/json' \    header "Authorization: Bearer <token>"</token></oml></pre>
View results	SELECT * FROM OML\$8017EE54_C305_47A5_9603_0EFA84B6B624_DATA_BIAS_JOB_RESULT
using SQL	ORDER BY SENSITIVE_FEATURE, STRATA_FEATURE

# **Objective:** *enable taking informed action*



## **Oracle Machine Learning family of components**

OML Component	Autonomous Database Serverless   DR	Autonomous Database Dedicated   C@C	Oracle Database on premises, BDBS, CS, CI, C@C
OML4SQL API Build ML models and score data with no data movement using SQL and PL/SQL	<ul> <li></li> </ul>	<ul> <li></li> </ul>	<ul> <li></li> </ul>
<b>OML4Py API</b> Leverage the database as a high-performance compute engine from Python with in-database ML and Spatial AI	<ul> <li></li> </ul>		<
<b>OML4R API</b> Leverage the database as a high-performance compute engine from R with in-database ML			<ul> <li></li> </ul>
<b>OML Notebooks</b> SQL, PL/SQL, Python, R, conda, and markdown interpreters	<ul> <li></li> </ul>		
OML AutoML UI No-code automated modeling interface	×		
<b>OML Monitoring UIs</b> No-code user interface for monitoring changes in data and in- database ML model quality			
<b>OML Services</b> RESTful model management, deployment, monitoring, bias detection			
Oracle Data Miner SQL Developer extension with a drag-n-drop interface for creating ML methodologies	<ul> <li></li> </ul>		×

## For more information...

OML Webpage https://oracle.com/machine-learning

OML Blog https://bit.ly/omlblogs

OML GitHub Repository https://bit.ly/omlgithub

OML Office Hours https://bit.ly/omlofficehours

### Try on Oracle LiveLabs

Overview: https://bit.ly/omlfundamentalshol OML4Py: https://bit.ly/oml4pyhol All OML: https://bit.ly/omllivelabs

### **OML** Documentation

https://docs.oracle.com/en/database/oracle/machine-learning/oml-monitoring https://docs.oracle.com/en/database/oracle/machine-learning

Copyright © 2025 Oracle and/or its affiliates.



# Thank you

### Mark Hornick

mark.hornick@oracle.com MarkHornick Group: Oracle Machine Learning

Copyright © 2025 Oracle and/or its affiliates.



# Helpful Links –

ORACLE ANALYTICS VIDEOS: <a href="https://www.youtube.com/@OracleAnalytics/videos">https://www.youtube.com/@OracleAnalytics/videos</a>

OAC SEPTEMBER NEW FEATURES VIDEOS BY ORACLE: https://bit.ly/OACSept24Features

OAC NEW FEATURES DOCUMENTATION BY ORACLE:

https://docs.oracle.com/en/cloud/paas/analytics-cloud/acswn/index.html#GUID-CFF90F44-BCEB-49EE-B40B-8D040F02D476

**ORACLE ANALYTICS COMMUNITY:** https://community.oracle.com/products/oracleanalytics

ORACLE ANALYTICS LIBRARY/EXAMPLES:

https://www.oracle.com/business-analytics/data-visualization/examples/

ORACLE ANALYTICS LIVE DEMOS: https://www.oracle.com/business-analytics/data-visualization/demos/



# Future & Past TechCasts:



Data and Model Monitoring – The Step Not To Skip In Solution Deployment

Presented by Mark Hornick

May 29th

Thwart Toil Through Tiles: Leveraging Oracle 23ai's Latest Geospatial Features

Presented by Jim Czuprynski

## TechCast Archive

2025		2024	2023	2022	2021		2020
Date	Title			Presenter(s)		Replay	Download(s)
Apr 3	Exploring Relationships in Your Data With Oracle Analytic Cloud (OAC)			Melli Annamalai, Philippe Lions & Gautam Pisharam		Video	Slides
Mar 6	Automating Oracle Analytics Cloud Administration with REST APIs			Joel Acha		Video	Slides
Feb 6	Our Favorite New Features in OAC: February 2024 and January 2025 Releases			Dan Vlamis, Wayne Van Sluys, Philippe Lions	Gautam Pisharam,	Video	Slides
Jan 23	Leveraging Vector Search for RAG in Generative AI		Kai Yu		Video	Slides	
Jan 9	The Oracle AI Microservices Sandbox for RAG Rapid Prototyping		Corrado De Bari, Mark Nelson, & John Lathouwers		Video	Slides	

Submit a topic to share at https://andouc.org/techcasts/



www.andouc.org

