ORACLE

Art of Possible with AI & Data Science

Healthcare Did Not Attend (DNA) Risk Model

15/09/23

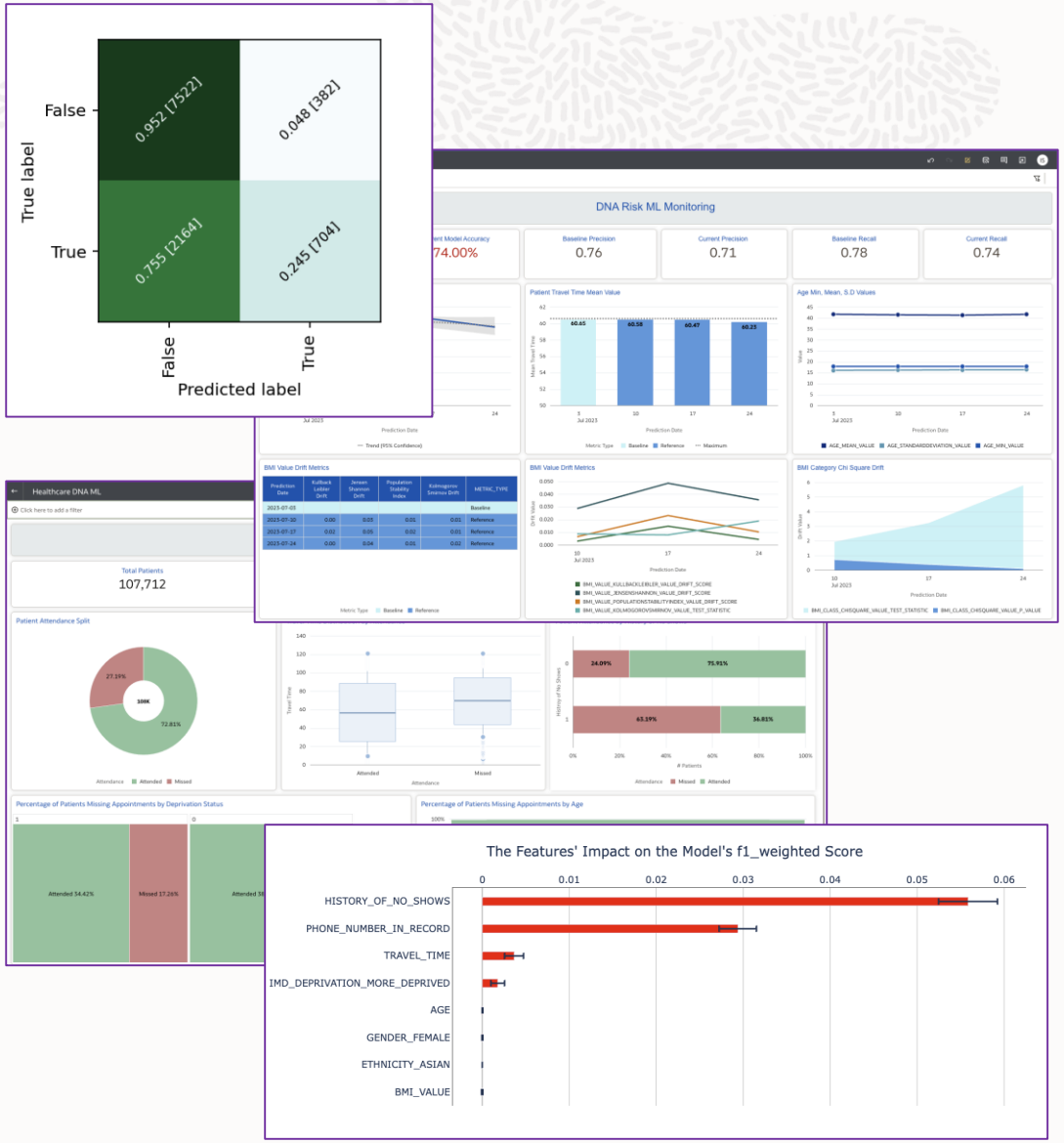
EMEA Data Science Team

Demo Flow

1. Demo Inspiration & Target Personas

2. How we achieved this?
   - Data Preparation & Exploratory Analysis
   - Citizen Data Science: ML Model Prototyping
   - DNA Risk Modelling, Evaluation & Explainability
   - Actionable Insights for Business
   - ML Operations

3. Behind the Scenes
   - Oracle Data Platform



Demo Inspiration & Target Personas

Demo Inspiration

This demo is going to showcase Oracle Data Science capabilities to predict the likelihood of a patient not attending their scheduled appointment.

The objective it not only to make accurate predictions, but to also understand the reasons why our Machine Learning Model has predicted each patient to be a show or a no show.

By exploring explainability of a model we will be able to answer the question of 'why' a patient will likely be a no show so that the practice operational team can follow up with a friendly reminder or reschedule for a more convenient time.

*"According to Hospital Episode Statistics (HES) data in 2019-20, outpatient attendances as a proportion of total appointments have decreased from 80.1% to 77.2% in the last 10 years. NHS England and Improvement says DNAs have an enormous impact on the healthcare system in terms of increasing both costs and waiting times." **

*"Of the 103 million outpatient appointments booked in 2021/22, 7.6% ended in a 'Did Not Attend'; this equates to an average of 650,000 monthly appointment slots." ***

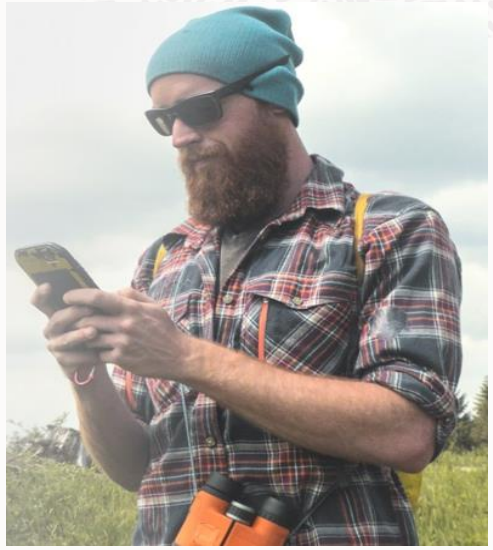
* https://econsult.net/blog/how-much-money-could-the-nhs-save-by-reducing-outpatient-dnas

** https://www.england.nhs.uk/long-read/reducing-did-not-attends-dnas-in-outpatient-services

Target Personas

Practice Operational Manager would like to understand the list of high-risk patients for not attending their scheduled appointment in the upcoming week along with the top reasons why, in order for the team to send reminders or to help reschedule appointments.

Citizen Data Scientist within the Operational Team would like to gain insights on why patients don't turn up to their appointments. She would like to experiment and test if building a prototype ML Model can help to predict which patients will not show up to their appointment.



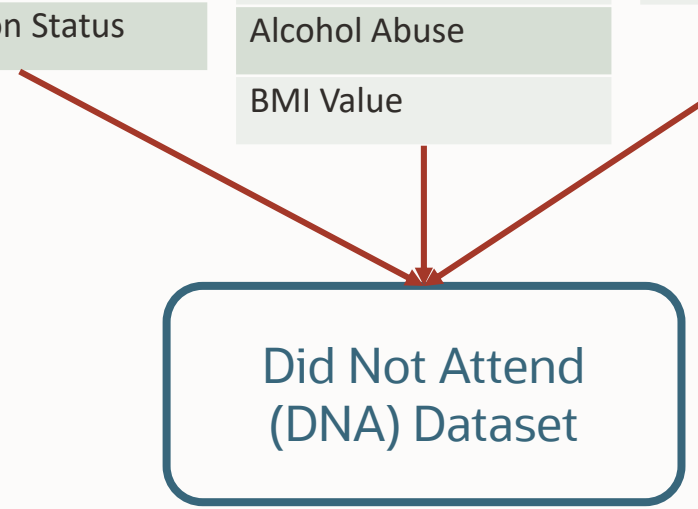
Data Scientist builds on the work done by the Citizen Data Scientist to build a more robust DNA Risk Model which is able to not only predict high-risk patients that will be a no-show, but also provide reasons why. Data Scientist also wants to track the Models performance over time.

How have we achieved this?

Data Preparation & Exploratory Analysis

As a Citizen Data Scientist I want to be able to explore and gain insights on my data to start understanding any data quality issues, trends and patterns in my data that can help identify reasons why patients miss appointments.

Demographic Data	Health Data	Behavioral Data
Patient ID	Anxiety	Travel Time
Age	Mental Illness	Phone Number in Record
Gender	Sleep Disorder	History of No Shows
Ethnicity	Antidepressants	Did Not Attend
Deprivation Status	Alcohol Abuse	
	BMI Value	



Click here to add a filter

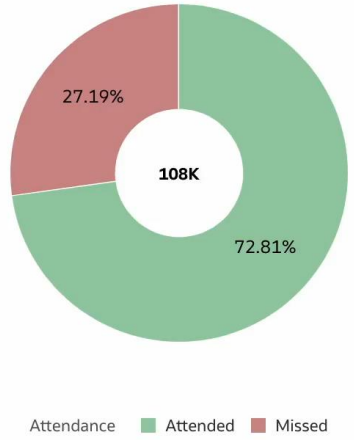
DNA Risk Exploratory Analysis

Total Patients
107,712

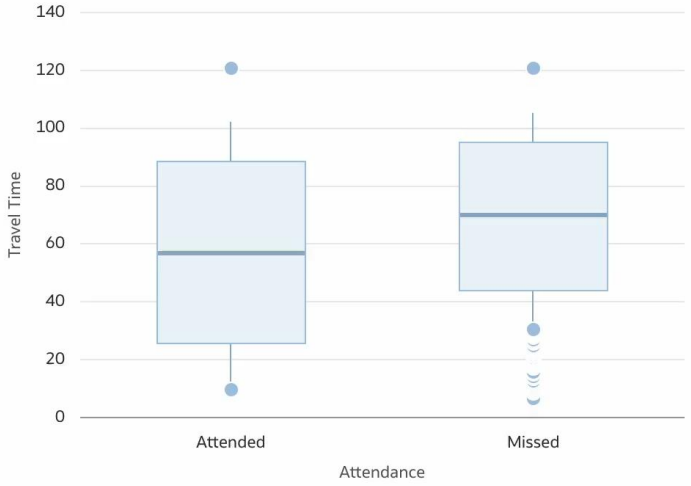
Attended Appointments
78,420

Missed Appointments
29,292

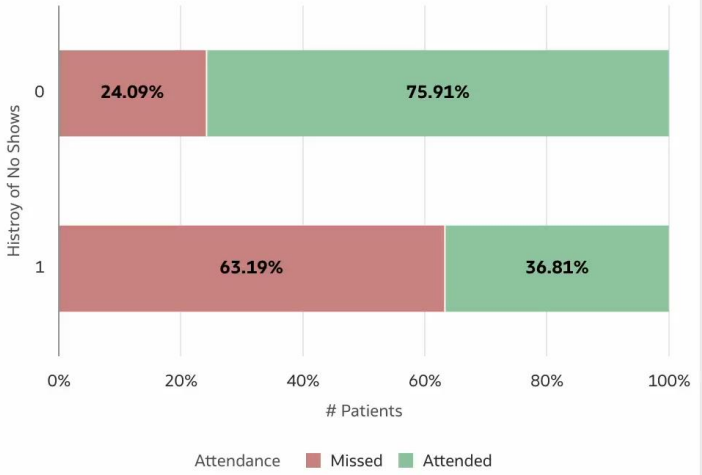
Patient Attendance Split



Travel Time Distribution by Attendance

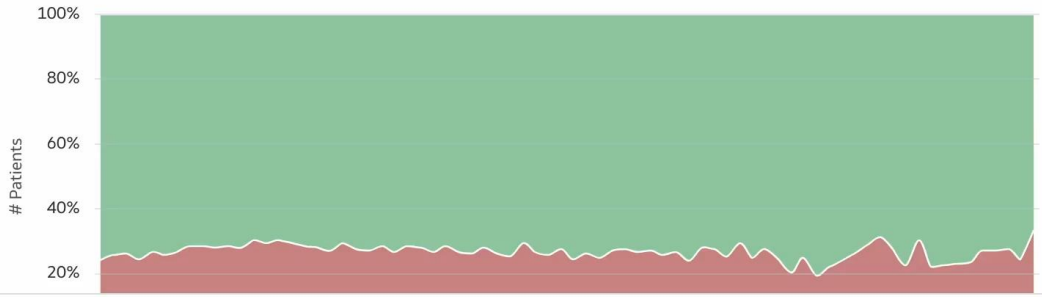


Patient Attendance by History of No Shows



Percentage of Patients Missing Appointments by Deprivation Status

Percentage of Patients Missing Appointments by Age



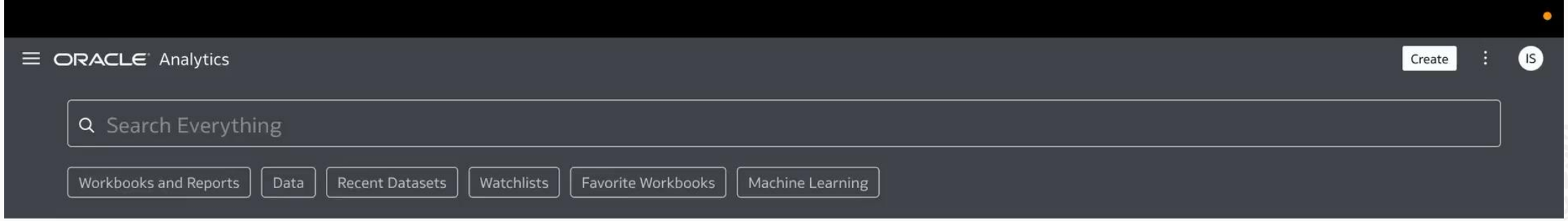
Citizen Data Science: Prototyping DNA Risk Model

As a Citizen Data Scientist sitting within my LoB, I want to be able to build a prototype ML Model to evaluate whether ML can predict which patients are likely to be a no-show.

Who is a Citizen Data Scientist?

Citizen Data Scientists typically sit within LOB's and bridge the knowledge gap between Data Analysts and Data Scientists. They work closely with both while possessing deep business domain knowledge and technical skills. The main job function of a Citizen Data Scientist is not to build Machine Learning Models.

Citizen Data Scientists often are not proficient in coding languages so prefer to interact with drag and drop tools to help analyse data, extract insights and build up predictive models.



Workbooks View all

Healthcare DNA ML
ismail.s.syed@oracle.com

NHSE GP Reviews
ismail.s.syed@oracle.com

Officer Availability
ismail.s.syed@oracle.com

Footfall Forecasting
ismail.s.syed@oracle.com

Mapping
ismail.s.syed@oracle.com

V1
aydin.heydari@oracle.com

AI-Vision-CarParking-o...
ismail.s.syed@oracle.com

Datasets View all

DNA ML Dataset | DNA ML Predictions | DNA ML Monitoring | BC_Demo_File | NHSE GP Reviews | AIA_Footfall_Fore... | AIA_Prophet_Fore... | AIA_Carpark_Utili...

Data Flows View all

DF_AIVISION_Obj... | DF_Sample_Order... | HC_Suppiler Spen... | HC_Teammate Co... | Recruitment | DF_ML_Predict_A... | DF_ML_Training_... | Naive Bayes - Attri...

Machine Learning

Pretrained Object ... | ML_Predict_Attriti... | Breast_Cancer_Tr...

DNA Risk Model Development

After discussing this use cases with our Citizen Data Scientist and Practice Operational Manager, I will build a DNA Risk Model to predict which patients are likely to be a no-show along with explainability rules.

What is a DNA Risk Model?

Using historic data of previous no-shows of patients we will train a Machine Learning Model to predict which patients are likely to be a no-show in the future with a degree of confidence (risk score).

What is Model Explainability?

Model Explainability can fall into two categories:

- Global Explainability describes the overall behaviour of a ML Model. Such as what are the most influential features for predicting our target variable.

- Local Explainability explains the rules (decisions) behind why the model made a certain prediction for an individual sample or in our case a patient.

ORACLE Cloud DSNS_ISYED Session remaining: 5hr 35min Extend Sign Out

File Edit View Run Kernel Git Tabs Settings Help

/ ... / AI_Approach / DNA-Risk-Model /

Name	Last Modified
input_data	12 days ago
models	4 days ago
output_data	12 days ago
training-test-data	4 days ago
00-dna-risk-feature...	12 days ago
01-dna-risk-model-...	2 minutes ago
02-dna-risk-model-...	an hour ago
03-dna-risk-ml-insi...	12 days ago
banner.png	12 days ago
mlm_insights-0.1.1_...	12 days ago

Launcher | 01-dna-risk-model-build.ip | 02-dna-risk-model-predict

Markdown Python [conda env:automlx_p38_cpu_v4]

Oracle AutoML & Explainability: Patient Do Not Attend Risk Model

Notebook Description: In this Notebook we will look at how you can utilise Oracle's low code AutoMLx and ADS Python Libraries to address the Data Science Lifecyle, from Data Exploration using ADS, Model Training & Tuning using AutoML, Model Evaluation using ADSEvaluator, AutoMLx for both Global and Local Explainability along with using ADS to save the Model to the centralised Model Catalog.

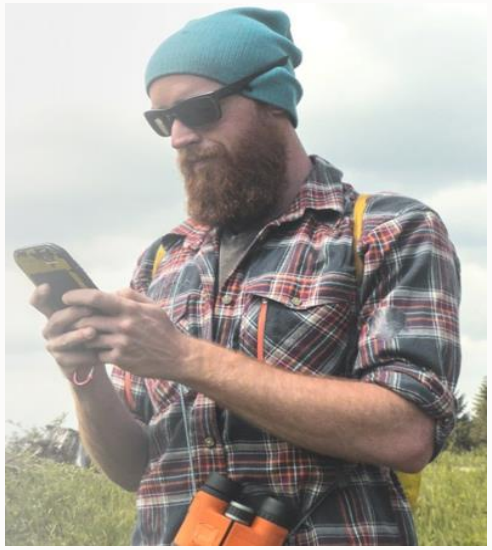
In this use case we will utilise a dummy dataset to predict and associate a risk score on whether or not a patient is likely to show or be a no show for an upcoming appointment based on various demographic and behavioural features of the patient.

Imports

[1]:
```
import os
import numpy as np
import pandas as pd
from dotenv import load_dotenv
import oci
import ads
import automl
from automl import init
from ads.evaluations.evaluator import ADSEvaluator
from ads.common.data import ADSData
from ads.common.model import ADSModel
from ads.model.generic_model import GenericModel
from ads.common.model_metadata import UseCaseType
```

0 3 Python [conda env:automlx_p38_cpu_v4] | Idle Saving completed Mode: Command Ln 1, Col 1 01-dna-risk-model-build.ipynb

Actionable Insights for Business



Data Scientist is able to provide a dataset containing predictions made from the DNA Risk Model on the patients likely to be a no-show along with explainability rules.

Citizen Data Scientist is able to present the predictions and rules provided by the Data Scientist in a business friendly dashboard which the Practice Operational Team can interact with.

Practice Operational Manager can use the interactive dashboard to identify the patients at most risk of being at being a no-show for the upcoming week and send reminders or help put preventative measures in place depending on the explainability rules.

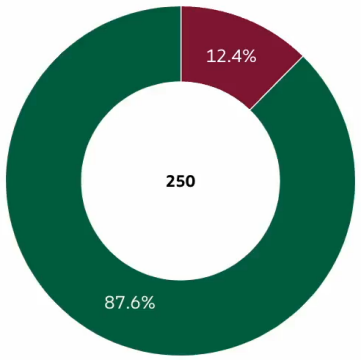
Healthcare DNA ML

Click here to add a filter

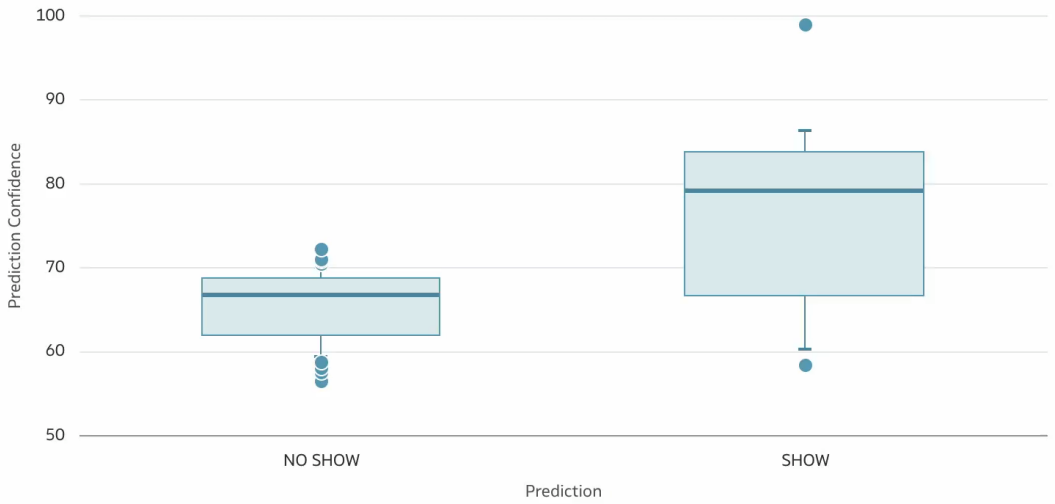
DNA Risk Prediction Results

Predictions Made	Predicted Show	Actual Show	Predicted No Show	Actual Show
250	219	181	31	69

Predicted Attendance Distribution



Prediction Confidence Distribution



Patient No Show List

Patient ID	Phone Number in Record	Travel Time	Prediction	Confidence	Explainability Rule 1	Explainability Rule 2	Explainability Rule 3	Explainability Rule 4	Explainability Rule 5
Synth103785	1	83.85	NO SHOW	62.06	HISTORY_OF_NO_SHOWS = 1	TRAVEL_TIME = 83.84878885	IMD_DEPRIVATION_MORE_DEPRIVED = 0	PHONE_NUMBER_IN_RECORD = 1	AGE = 31
Synth116588	0	106.84	NO SHOW	60.64	PHONE_NUMBER_IN_RECORD = 0	IMD_DEPRIVATION_MORE_DEPRIVED = 0	TRAVEL_TIME = 106.8375889	HISTORY_OF_NO_SHOWS = 0	AGE = 54
Synth122455	1	92.85	NO SHOW	68.09	HISTORY_OF_NO_SHOWS = 1	IMD_DEPRIVATION_MORE_DEPRIVED = 1	TRAVEL_TIME = 92.8536566	PHONE_NUMBER_IN_RECORD = 1	AGE = 20

Exploratory Analysis | **Prediction Results** | ML Monitoring

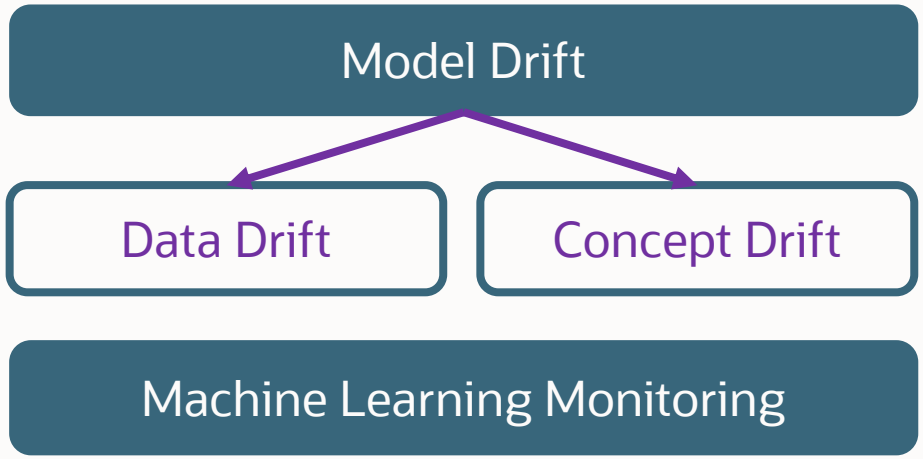
MLOps: Machine Learning Monitoring

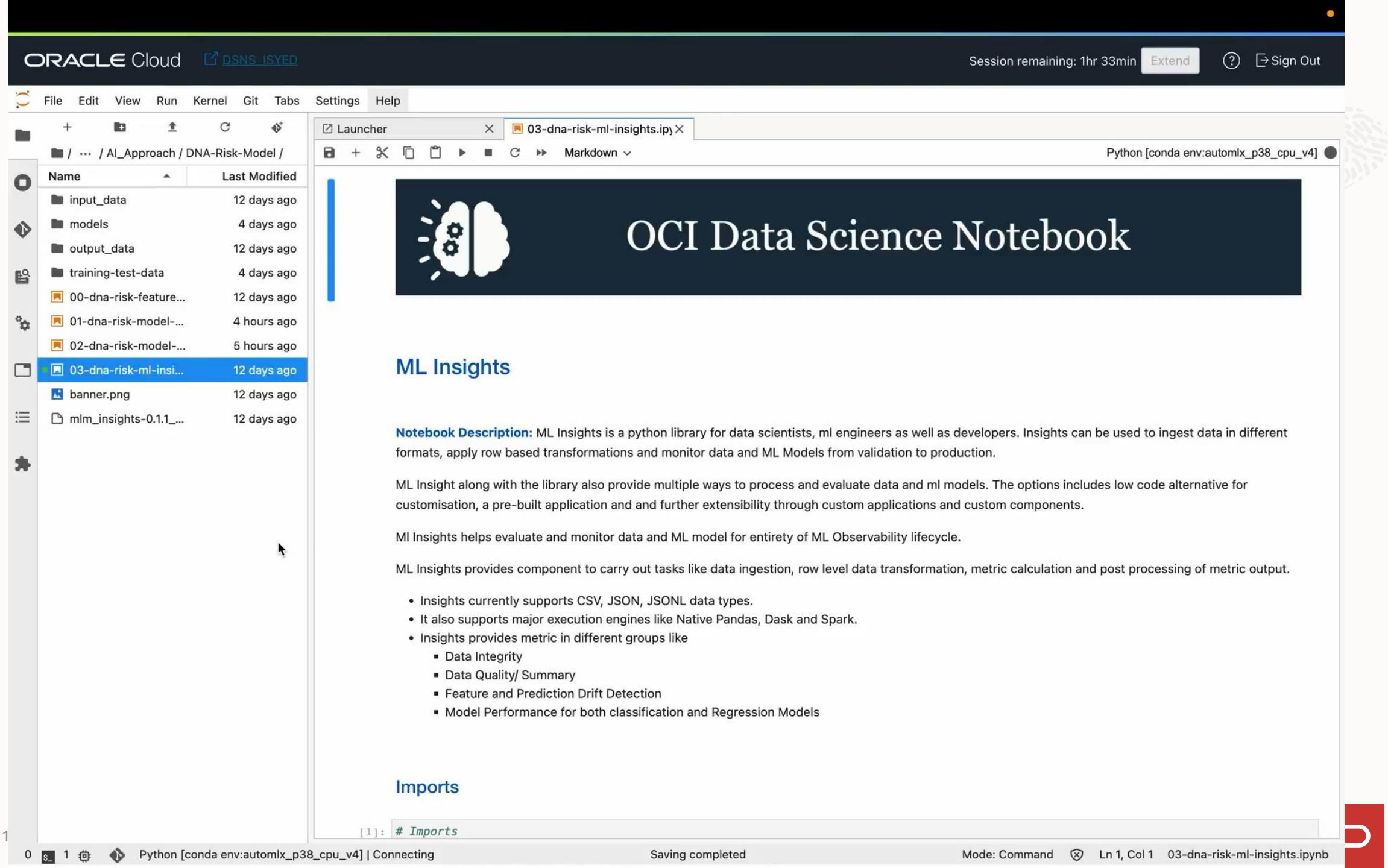
Now that the DNA Risk Model is in Production, I must now keep an eye on the Model Performance and take action on any Data or Model Drift detected.

What is Machine Learning Monitoring?

Machine Learning Model are most effective when first in production and tend to degrade over time.

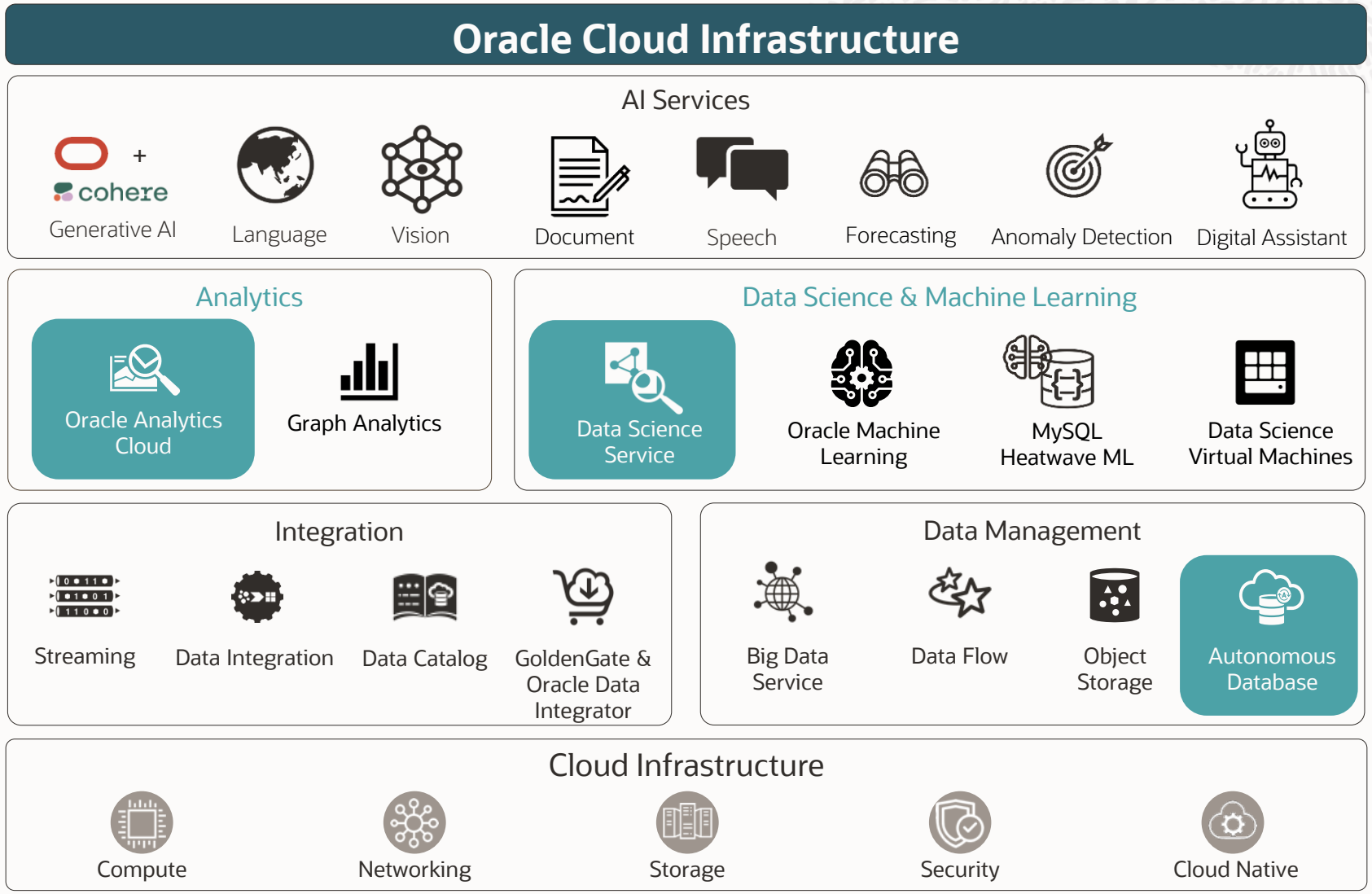
ML Monitoring allows you to keep track of your data and model's performance over time and identify when certain metrics drops below an acceptability threshold which might be an indication of needing to retrain your model on new data.



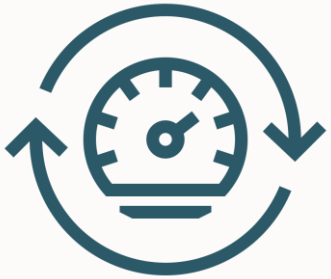

ORACLE Cloud
DSNS_ISYED
Session remaining: 1hr 33min
Extend
Sign Out
File Edit View Run Kernel Git Tabs Settings Help
/ ... / AI_Approach / DNA-Risk-Model /
Name
Last Modified
input_data 12 days ago
models 4 days ago
output_data 12 days ago
training-test-data 4 days ago
00-dna-risk-feature... 12 days ago
01-dna-risk-model-... 4 hours ago
02-dna-risk-model-... 5 hours ago
03-dna-risk-ml-insi... 12 days ago
banner.png 12 days ago
mlm_insights-0.1.1_... 12 days ago
Launcher
03-dna-risk-ml-insights.ipy
Markdown
Python [conda env:automlx_p38_cpu_v4]
OCI Data Science Notebook
ML Insights
Notebook Description: ML Insights is a python library for data scientists, ml engineers as well as developers. Insights can be used to ingest data in different formats, apply row based transformations and monitor data and ML Models from validation to production.
ML Insight along with the library also provide multiple ways to process and evaluate data and ml models. The options includes low code alternative for customisation, a pre-built application and and further extensibility through custom applications and custom components.
Ml Insights helps evaluate and monitor data and ML model for entirety of ML Observability lifecycle.
ML Insights provides component to carry out tasks like data ingestion, row level data transformation, metric calculation and post processing of metric output.
Insights currently supports CSV, JSON, JSONL data types.
It also supports major execution engines like Native Pandas, Dask and Spark.
Insights provides metric in different groups like
Data Integrity
Data Quality/ Summary
Feature and Prediction Drift Detection
Model Performance for both classification and Regression Models
Imports
[1]: # Imports
0 1 Python [conda env:automlx_p38_cpu_v4] | Connecting
Saving completed
Mode: Command
Ln 1, Col 1
03-dna-risk-ml-insights.ipynb

Behind the Scenes

Behind the Scenes



Autonomous Database



Self-Driving

- Scale-out database with fault-tolerance and DR
- Runs on enterprise-proven Exadata platform
- Full compatibility with existing enterprise databases

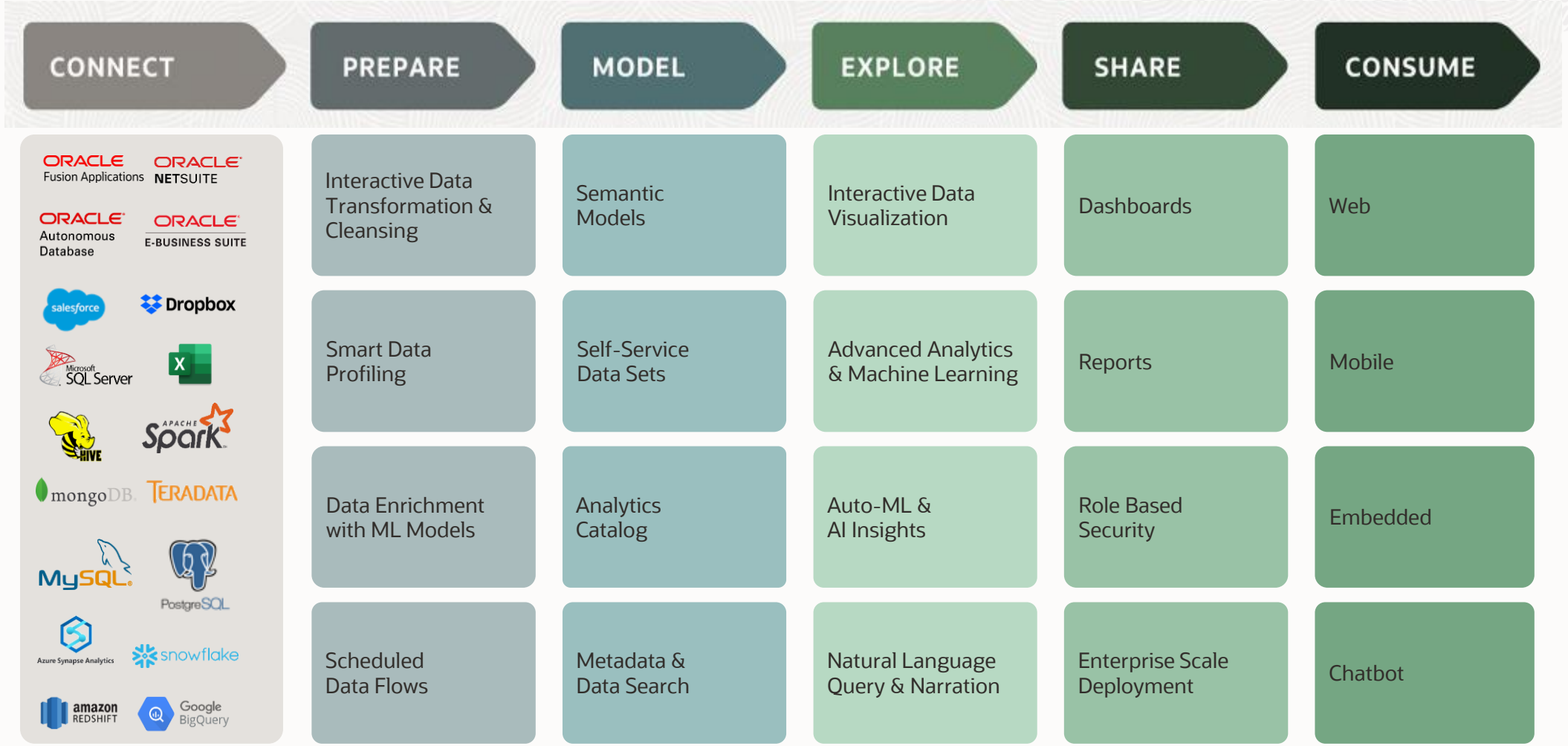
Self-Securing

- Automatically applies security updates online
- Secure configuration with full database encryption
- Sensitive data hidden from Oracle or customer admins

Self-Repairing

- Recovers automatically from any failure
- 99.995% uptime including maintenance
- Elastically scales compute or storage as needed

Oracle Analytics Cloud



Oracle Cloud Infrastructure Data Science

Support for Python and Open Source

Accelerated

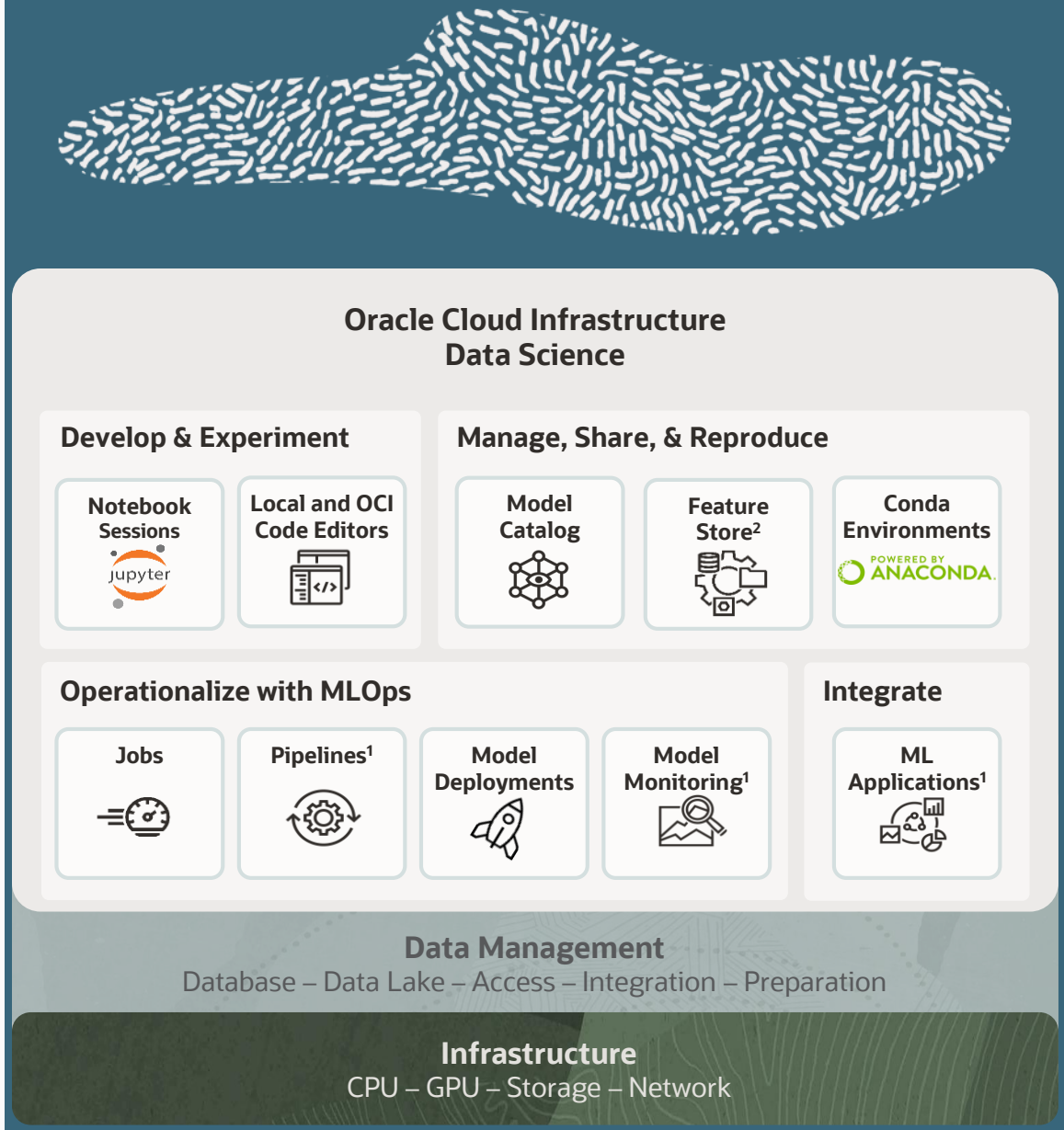
Allow data scientists to work the way they want to, and provide access to automated workflows, the best of open source, and a streamlined approach to building models.

Collaborative

Enable data science teams to work together with ways to share and reproduce models in a structured, secure way for enterprise-grade results.

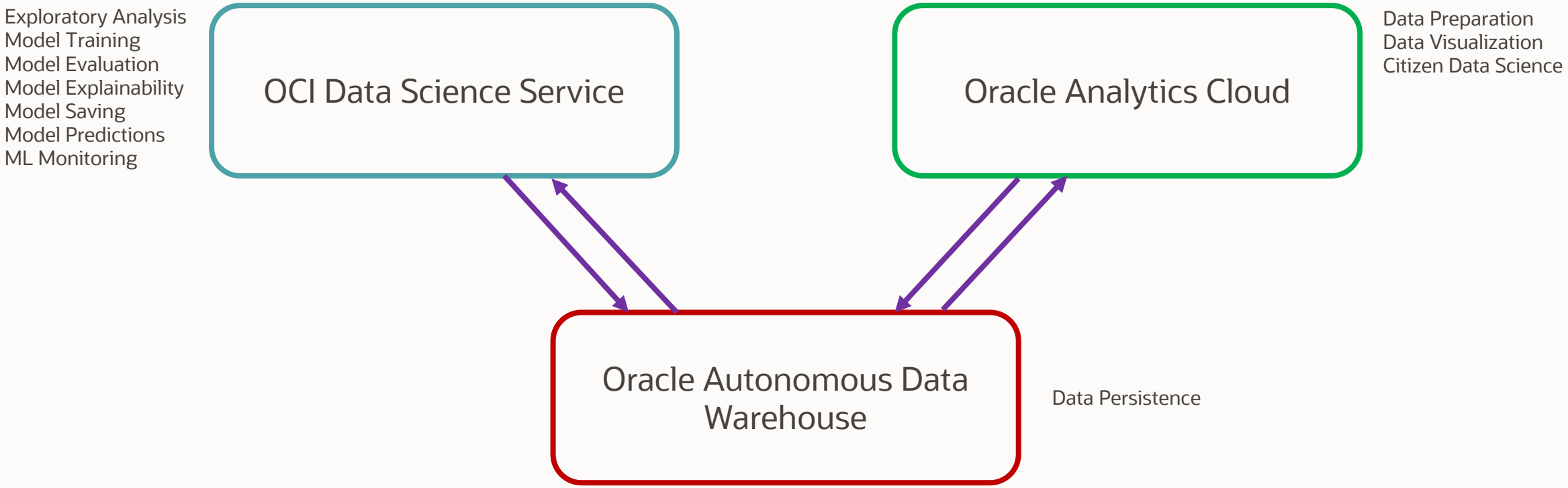
Enterprise-Grade

Provide a fully managed platform built to meet the needs of the modern enterprise.

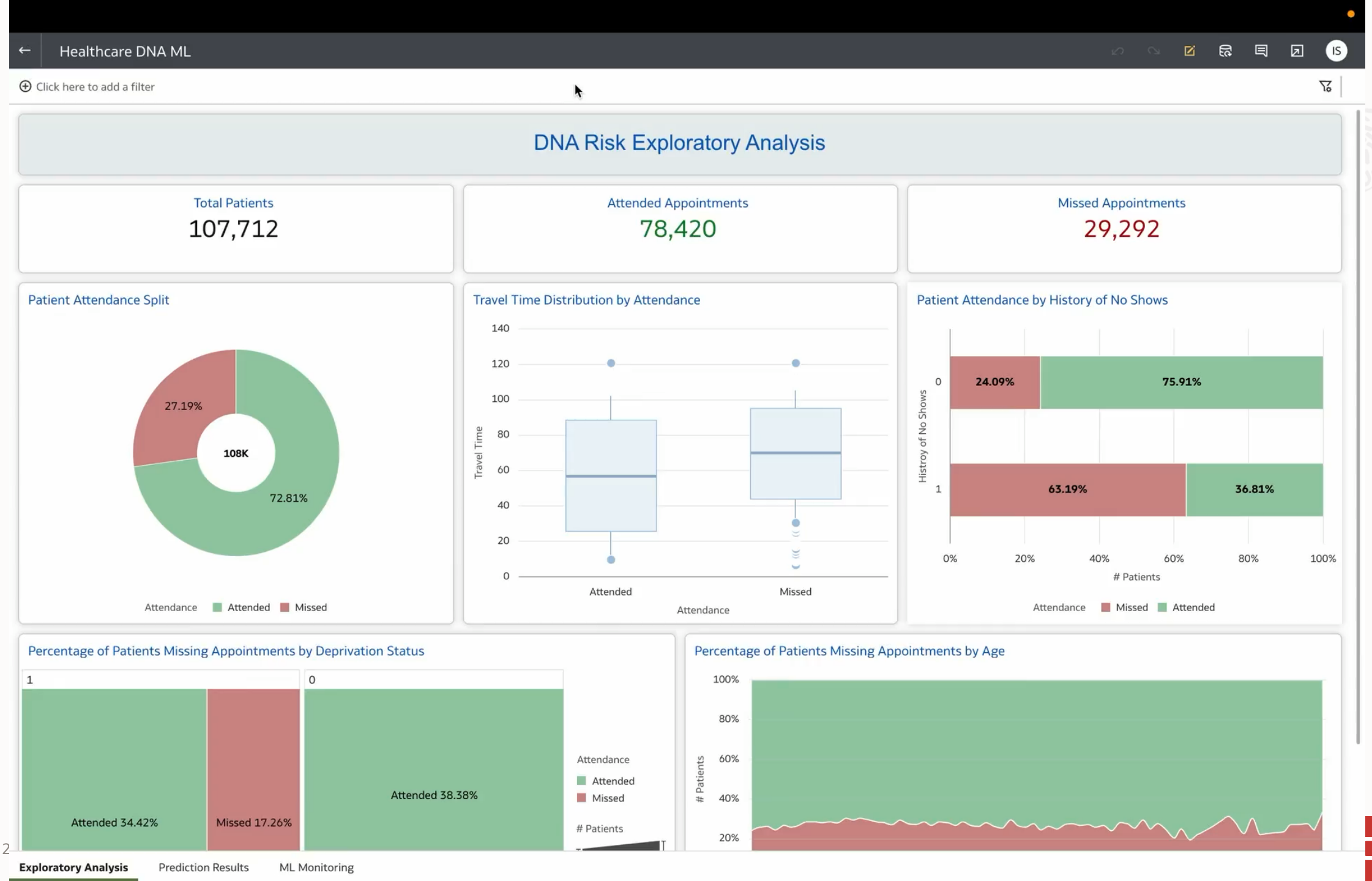


[1] In preview
[2] Coming soon

High Level Architecture



Full Demo Recording


Healthcare DNA ML
IS
Click here to add a filter
DNA Risk Exploratory Analysis
Total Patients
107,712
Attended Appointments
78,420
Missed Appointments
29,292
Patient Attendance Split
27.19%
108K
72.81%
Attendance
Attended
Missed
Travel Time Distribution by Attendance
140
120
100
80
60
40
20
0
Travel Time
Attended
Missed
Attendance
Patient Attendance by History of No Shows
Histroy of No Shows
0
1
24.09%
75.91%
63.19%
36.81%
0%
20%
40%
60%
80%
100%
Patients
Attendance
Missed
Attended
Percentage of Patients Missing Appointments by Deprivation Status
1
0
Attended 34.42%
Missed 17.26%
Attended 38.38%
Attendance
Attended
Missed
Patients
Percentage of Patients Missing Appointments by Age
100%
80%
60%
40%
20%
Patients
Exploratory Analysis
Prediction Results
ML Monitoring

Thank you

EMEA Data Science Team


ORACLE