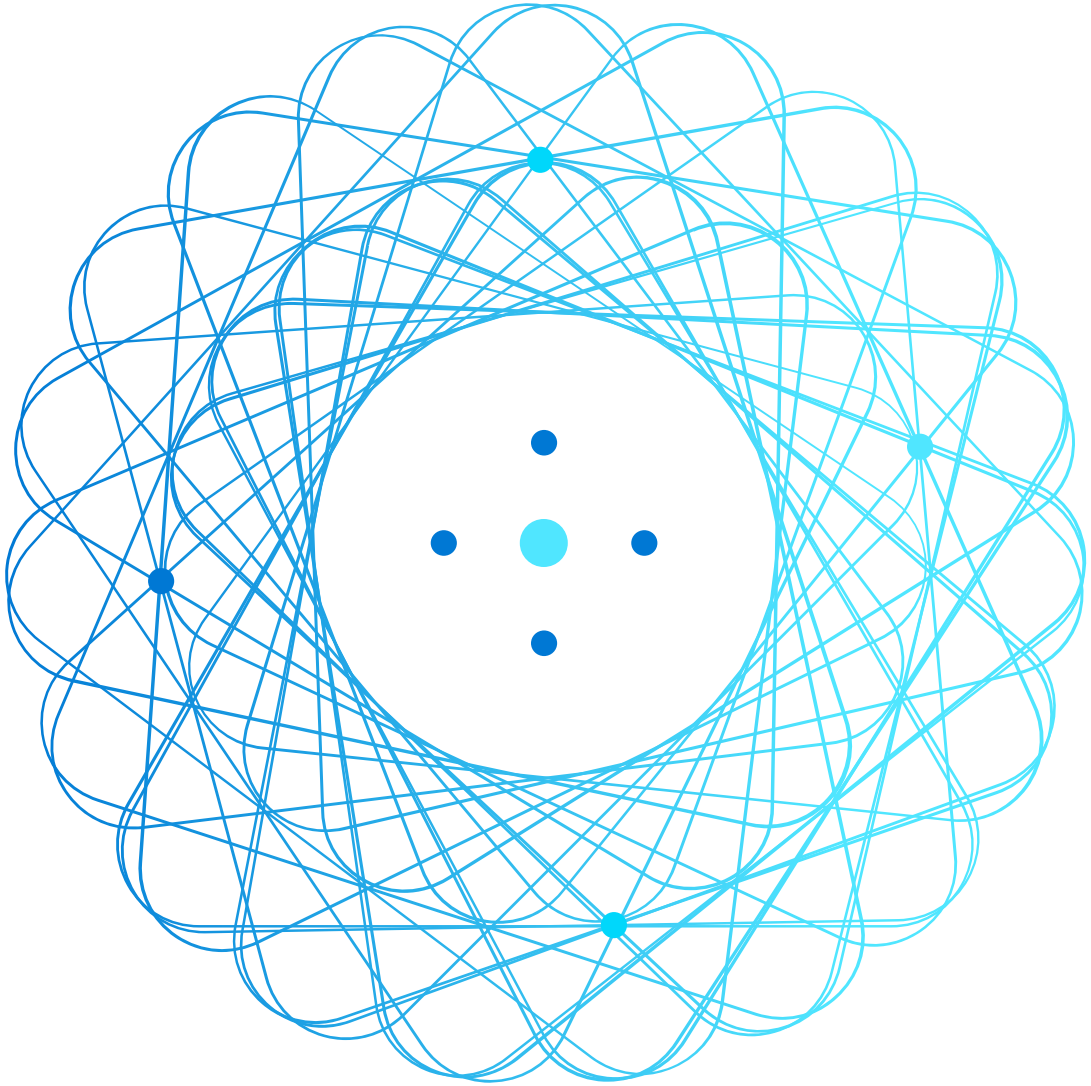
Microsoft Azure

Module 10: Monitoring Models

Mohammed Arif

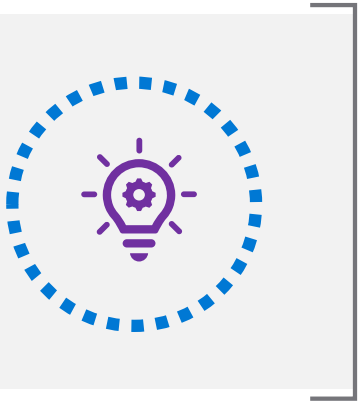


Agenda

Monitoring Models with Application Insights

Monitoring Data Drift

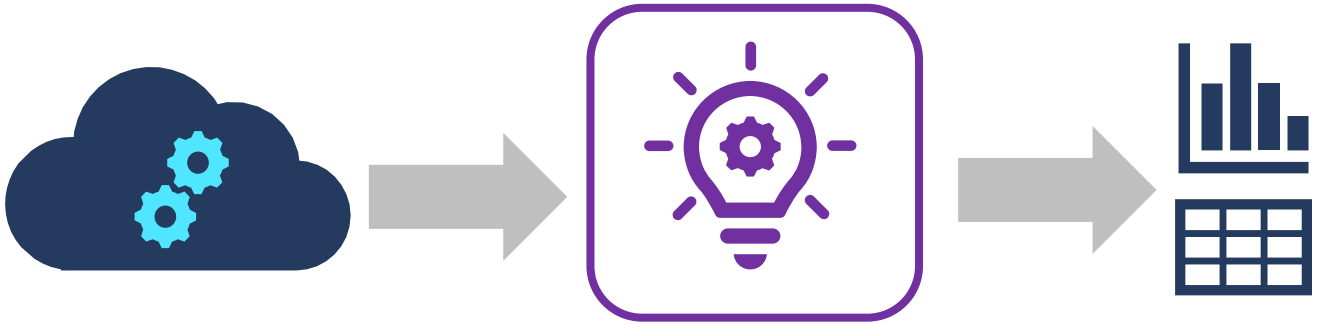
Monitoring Models with Application Insights



What is Application Insights?

An Application Performance Management service in Azure

Enables capture, storage, and analysis of telemetry data



Enabling Application Insights

Determine the Application Insights resource for your workspace

```
ws.get_details()['applicationInsights']
```

Enable in a new service deployment configuration using the SDK:

```
deploy_config = Webservice.deploy_configuration(enable_app_insights=True)
```

Enable for existing deployed services:

Configure AKS deployment in Azure Machine Learning studio

Update deployed service using the SDK

```
service.update(enable_app_insights=True)
```

Capturing and Viewing Application Insights Data

Print log data in the scoring script

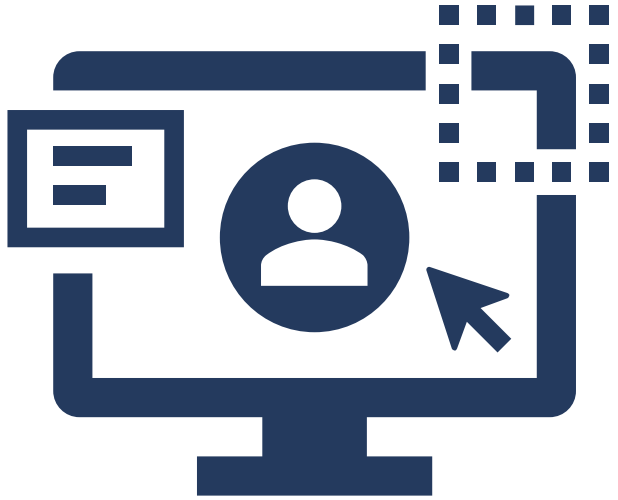
```
def init():
    model = joblib.load(Model.get_model_path('my_model'))
def run(raw_data):
    data = json.loads(raw_data)['data']
    predictions = model.predict(data)
    log_txt = 'Data:' + str(data) + ' - Predictions:' + str(predictions)
    print(log_txt)
```

Query Logs in Application Insights

```
traces
|where  message == "STDOUT" and customDimensions.["Service Name"] = "my-svc"
| project  timestamp, customDimensions.Content
```

timestamp	customDimensions_Content
01/02/2020, 9:11:57.846 PM	Data:[[1, 2, 2.5, 3.1], [0, 1, 1,7, 2.1]] - Predictions:[0 1]

Lab: Monitor a Model

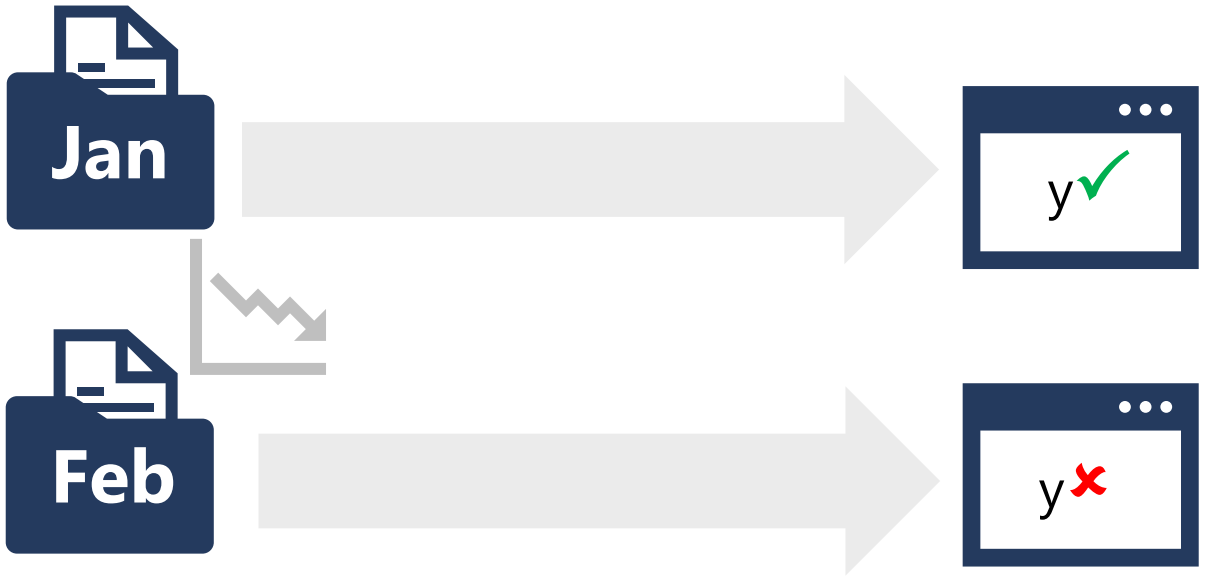


1. View the lab instructions at https://aka.ms/mslearn-dp100
2. Complete the **Monitor a model** exercise

Monitoring Data Drift

What is Data Drift?

Changing data trends that can affect the accuracy of trained models



Creating a Data Drift Monitor

Monitor by Comparing Datasets

Baseline dataset (original training data)

Target dataset for comparison over time (requires timestamp column)

Backfill to populate a data drift profile from target dataset

```
monitor = DataDriftDetector.create_from_datasets(ws, 'dataset-drift-detector',
                                                 baseline_data_set, target_data_set, ...)

backfill = monitor.backfill(dt.datetime.now() - dt.timedelta(days=30), dt.datetime.now())
```

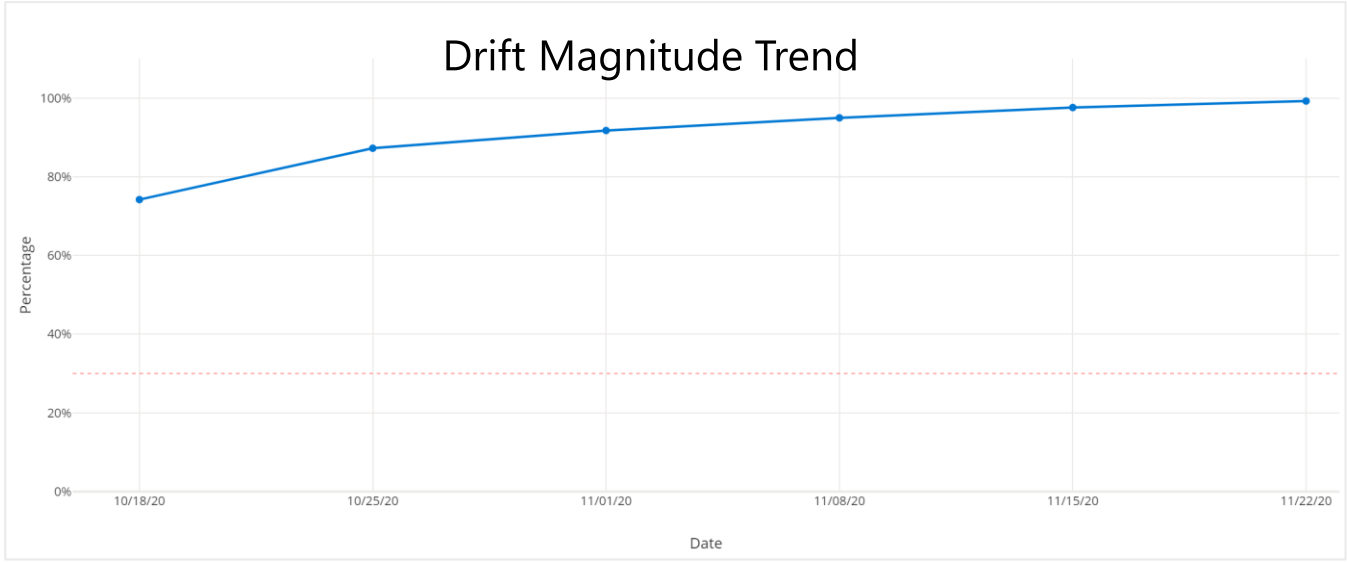
Data Drift Schedules and Alerts

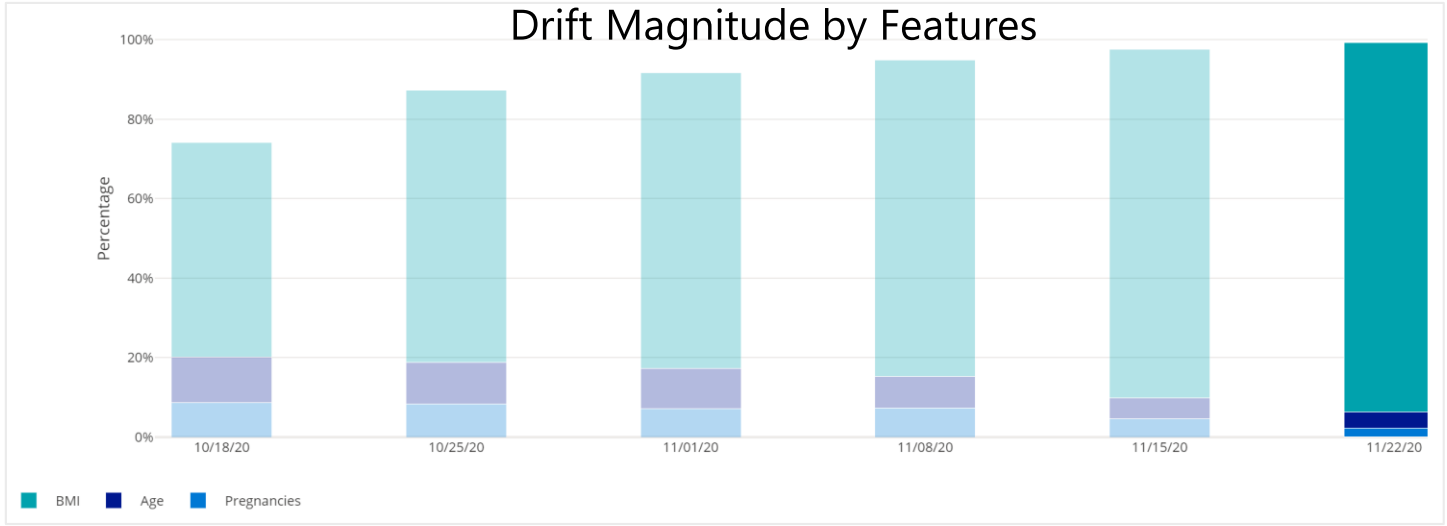
On creation, specify:

- Frequency
- Drift threshold for alerting
- Alert configuration
- Schedule start (for model data drift monitors)
- Data latency (for dataset data drift monitors)

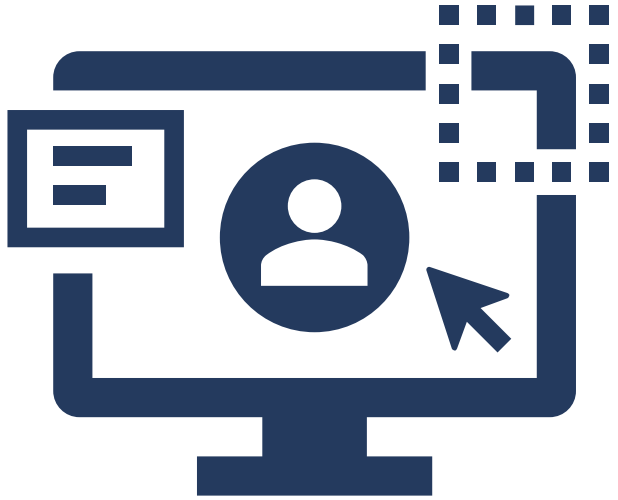
```
alert_email AlertConfiguration('data_scientists@contoso.com')
monitor = DataDriftDetector.create_from_datasets(ws, 'dataset-drift-detector',
                                                 baseline_data_set, target_data_set,
                                                 compute_target=cpu_cluster,
                                                 frequency='Week', latency=2,
                                                 drift_threshold=.3,
                                                 alert_configuration=alert_email)
```

Reviewing Data Drift





Lab: Monitor Data Drift



1. View the lab instructions at https://aka.ms/mslearn-dp100
2. Complete the **Monitor data drift** exercise

Knowledge check

You want to capture metrics from a real-time inference service and analyze them using Application Insights. What must you do in the scoring script for the service?

- ❑ Use the **Run.log** method to log the metrics.
- ❑ Save the metrics in the **./outputs** folder.
- ☑ Use a **print** statement to write the metrics in the STDOUT log.

You previously trained a model using a training dataset. You want to detect any data drift in the new data collected since the model was trained. What should you do?

- ❑ Create a new version of the dataset using only the new data; and retrain the model.
- ❑ Add the new data to the existing dataset and enable Application Insights for the service where the model is deployed.
- ☑ Create a new dataset using the new data and a timestamp column; and create a data drift monitor that uses the training dataset as a baseline and the new dataset as a target.

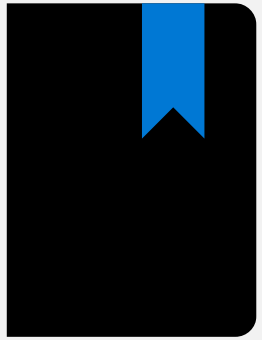
References

Microsoft Learn: Monitor models with Azure Machine Learning
https://docs.microsoft.com/learn/modules/monitor-models-with-azure-machine-learning

Microsoft Learn: Monitor data drift with Azure Machine Learning
https://docs.microsoft.com/learn/modules/monitor-data-drift-with-azure-machine-learning

Azure Machine Learning monitoring with Application Insights documentation
https://docs.microsoft.com/azure/machine-learning/how-to-enable-app-insights

Azure Machine Learning data drift documentation
https://docs.microsoft.com/azure/machine-learning/how-to-monitor-datasets



Microsoft Azure