

Amazon Lookout for Metrics ワークショップ

2021/05/21

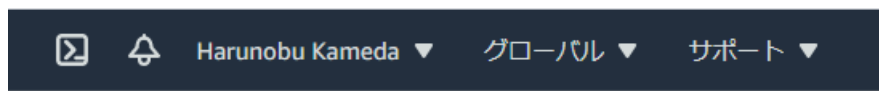
シニアエバンジェリスト 亀田

[S3 バケットの準備とデータのアップロード]

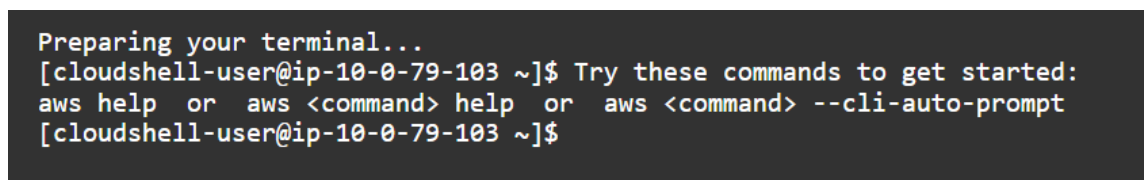
- 適当な名前で S3 バケットを作成し、その名前を手元のメモにコピーしておきます。設定は全てデフォルトで問題ありません。



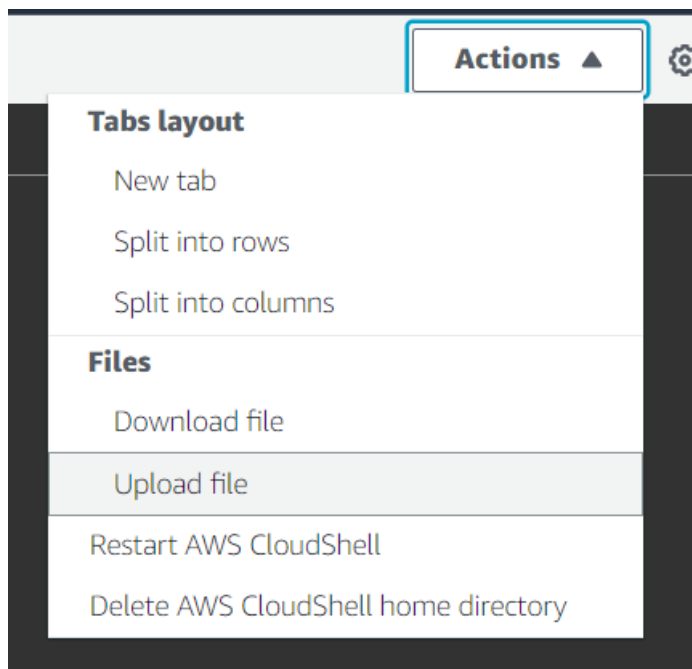
- Github のレポジトリから[ecommerce.zip]をダウンロードします
- CloudShell を起動します (AWS マネージメントコンソール右上です)



- 起動にしばらく時間がかかるので待つと以下が表示され使えるようになります。



- 画面右上の[Actions]から[Upload File]を選びます



- [ecommerce.zip]をアップロードします。念のため `ls -l` コマンドで確認しておきます。

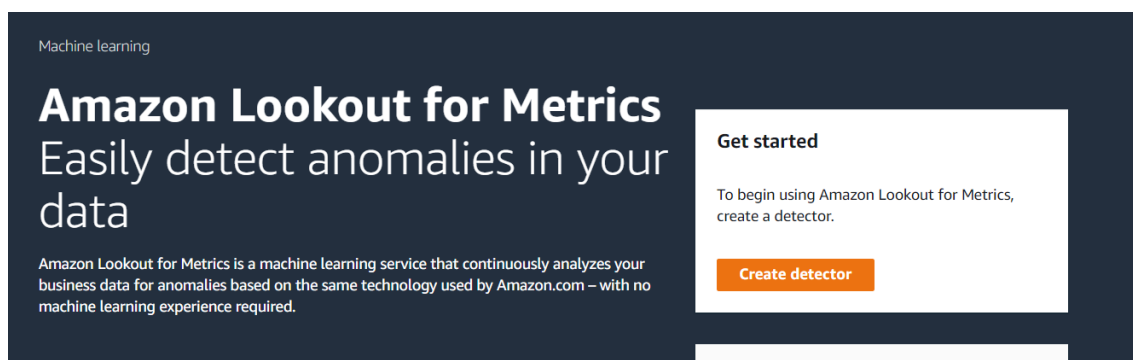
```
[cloudshell-user@ip-10-0-79-103 ~]$ ls -l
total 5956
-rw-rw-r-- 1 cloudshell-user cloudshell-user 6092322 May 18 04:50 ecommerce.zip
```

- [unzip ecommerce.zip]で解凍します
- [!aws s3 sync ./ecommerce/ s3://{s3_bucket}/ecommerce/]を実行し、作成されたディレクトリを纏めて S3 にアップロードします。(S3 マネージメントコンソールでもアップロード可能ですが数十倍高速です)。`{s3_bucket}`は先程作成した S3 バケット名に置き換えます
- アップロードが完了したら S3 バケットを見えます。Ecommerce フォルダの下に 1 つのファイルと 2 つのフォルダができていれば成功です。



[Lookout for Metrics で Detector（検出器）を作成]

- Management Console で Lookout for Metrics の画面に行き、[Create detector]を押します



- [Detector Name]に適切な名前を付け、[Interval]を[1 hour]に設定します。この Interval は皆さんが用意したデータ間隔と一致している必要があります

Detector name

detector20210520

The name can have up to 63 character max. Valid characters: a-z, A-Z, 0-9, - (hyphen) and _

Description - optional

My anomaly detector

The detector description can have up to 256 characters.

Interval

An interval is the amount of time between each analysis.

1 hour intervals

- [Create]を押します。

Detector20210520 [Info](#)

▼ How it works: Detector setup

Complete these steps to set up and activate your anomaly detector



Create a detector

A detector monitors your dataset, finds anomalies and analyzes their impact.



Add a dataset

Select the data that you want to monitor.

✔ Created

Add a dataset

- Detecor ができましたので、続いて[Add a dataset]を押して、学習準備を設定します
- 適度な名前を[Name]に設定します

Basic information

Name

dataset20210520

The dataset name must have 1 to 63 characters. Valid characters= a-z, A-Z, 0-9, - (hyphen) and _.

Description - *optional*

Description of my datasource and interval

The detector description can have up to 256 characters.

Timezone - *optional*

The timezone is used to interpret timestamps in your data.

UTC+00:00 (UTC, GMT, WET)

- [Datasource]に Amazon S3 を選択します

Datasource details

Read data from an Amazon S3 bucket that you manage or from a datasource that integrates with Amazon Lookout for Metrics.

To run the detector in backtest mode, use Amazon S3 for the datasource. Sample data is available in the developer guide. [Learn more](#)

Datasource Info

Choose the datasource that stores or generates your data.

Amazon S3

- [Detector Mode]を Backset に設定します

Detector mode Info

☒ Backtest

Find anomalies in historical data all at once.

☐ Continuous

Monitor live data to detect anomalies as they occur.

解説：Continuous モードはライブデータをもとに学習、そして異常検知を行います。
先程 S3 バケットにアップしたデータの live フォルダには以下のように 1 時間ごとに売り上げデータが命名規則に従い別ファイルとして生成されています。

Amazon S3 > lookoutformetrics0210520 > ecommerce/ > live/ > 20210101/ > 0000/

0000/

オブジェクト

プロパティ

オブジェクト (1)

オブジェクトは、Amazon S3 に保存された基本的なエンティティです。 [Amazon S3 インベントリ](#) を使用して、バケット内のすべてのオブジェクトにアクセスできるためには、明示的にアクセス権限を付与する必要があります。 [詳細はこちら](#)

URL をコピー

開く

ダウンロード

削除

アクション ▼

フォルダ

🔍

プレフィックスでオブジェクトを検索

<input type="checkbox"/>	名前 ▲	タイプ ▼	最終更新日時
<input type="checkbox"/>	20210101_000000.csv	csv	2021/05/20 03:09:59 PM JST

Continuous モードでは 1 時間間隔でこの中身を用いて、学習 & 推論を行います。皆さんの application は命名法則にしたがい設定した Interval 間隔でログを吐き出すようにしておく必要があります。命名法則にしたがう限り、設定不要で自動でファイルを読

み込み続けます。この際精度を向上させるために、[Historical Data]として過去のデータを学習させることも可能です。勿論この場合過去のデータと今の application へのリクエストにおいて、変更が加わっていないことが前提になります。

Historical data [Info](#)

Your continuous detector can learn from historical data to start detecting anomalies more quickly and with higher accuracy as soon as the detector is activated. If you do not provide historical data, your detector will learn from continuous data once it is activated, before it can start detecting anomalies.

☐ Use historical data

このハンズオンでは Continuous モードの場合、最低でも数時間待たないと結果が出力されないため、[Backtest]モードを使用します。

- [Historical data]に以下の値を入力します。{bucketname}は今日作成のものに置換してください

s3://{bucketname}/ecommerce/backtest/

ここで指定されたデータの 70%が学習に使用され、残り 30%が推論に使用されます。また設定した Interval の 285 回分のデータ（今回で言うと 285 時間分）が最低必要です。

- [Most recent date]に[2021/09/30]と設定します

Most recent date

Select the most recent date for which you have data in your source location e.g. if you have data from 10/1/2020 to 11/25/2020 then select 11/25/2020.

2021/09/30

< 2021年9月 >

日	月	火	水	木	金	土
29	30	31	1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	1	2

data from between 285 and 3000 intervals. The detector uses 70% of the data for
nt 30% for finding anomalies. Based on the end date, the backtest can use data
and 2021-01-01.

- [Permission]では、[Create a role]を選びます。現れたダイアログに S3 バケット名を入力し、[Create a role]を押します

Permissions

Service role [Info](#)

Amazon Lookout for Metrics needs permission to read data from your bucket. Choose a service role with access to your data, or create a role automatically.

Enter the ARN of a role ▲

Create a role

Enter the ARN of a role

Use an existing role

Create a service role×

Amazon Lookout for Metrics needs permission to access your data in Amazon S3. Choose a service role, or create a role in the IAM console with the AmazonPersonalizeFullAccess managed policy.

The service role you create grants Amazon Lookout for Metrics access to:

Amazon S3 buckets

lookoutformetrics0210520

Separate bucket names with commas.

CancelCreate a role

Permissions

Service role [Info](#)

Amazon Lookout for Metrics needs permission to read data from your bucket. Choose a service role with access to your data, or create a role automatically.

AmazonLookoutMetrics-ExecutionRole-1621492367008 ▼

✓

Successfully created a service role.
AmazonLookoutMetrics-ExecutionRole-1621492367008×

- [Next]を押します。(表示されるダイアログでエラーとなる場合、IAM 作成が間に合っていない可能性がありますので、もう一度 Next ボタンをおしてください)
- 次の画面の[Measures]と[Dimensions]を以下の画像のように設定します

Map fields [Info](#)

Measures [Info](#)

Measures are the key performance indicators (KPIs) that your detector tracks for anomalies.

Create measures using fields from your datasource. The detector will track these measures for anomalies. You can have up to 5 measures.

Field name

views ▼

Aggregate by

SUM ▼

revenue ▼

SUM ▼

Remove

Add measure

You can add up to 3 more measures

Dimensions - optional [Info](#)

Dimensions are categories that create segments in measure values. For example, by adding a region dimension you can organize your measure values into separate groups for each region.

Create dimensions using fields from your datasource. You can have up to 5 dimensions.

Dimensions

platform ▼

Remove

marketplace ▼

Remove

Add dimension

You can add up to 3 more dimensions.

解説：Measure は異常検知を行う対象となる値です。Dimensions はそれに影響を与える可能性のある項目です。例えば対象データが Global で展開されている EC サイトの売り上げ及びページビューとした場合、Dimensions には国やプラットフォーム（PC 向けやモバイル向け等）をセットすることで、検知された異常に影響を与えているパラメータを知ることができます

- Timestamp に[yyyy-MM-dd HH:mm:ss]を設定します。

Timestamp

The timestamp indicates when each data point was generated.

Specify a field from your datasource to use as a timestamp. You must have one timestamp.

Timestamp

timestamp ▼

Format

yyyy-MM-dd HH:mm:ss

yyyy-MM-dd HH:mm:ss (24 hour format)
 yyyy-MM-dd hh:mm:ss a (12 hour format)

- [Next]をおします。次に表示される確認画面で[Save and activate]を押します。さらに表示されるダイアログで[Active]を押すと、学習が開始されます

ⓘ The detector is learning.
 The detector uses data from several intervals to learn before finding anomalies. While it learns, you can configure alerts.

Add alert ×

[Amazon Lookout for Metrics](#) > [Detectors](#) > detector20210520

▼ How it works: Detector setup

Complete these steps to set up and activate your anomaly detector.

Create a detector

A detector monitors your dataset, finds anomalies and analyzes their impact.

✔ Created

Add a dataset

Select the data that you want to monitor.

✔ 1 added

Activate detector

Activate the detector to start monitoring the data for anomalies.

⋮ Activating backtest...

数十分かかるので待ちます

- しばらく待つと学習が完了しステータスが変更になります。



Activate detector

Activate the detector to start monitoring the data for anomalies.

✔ Backtest in progress

推論が開始され始めていますので、[View anomalies]を押すと検知された異常が憑依されています。

View and tune results.



View anomalies

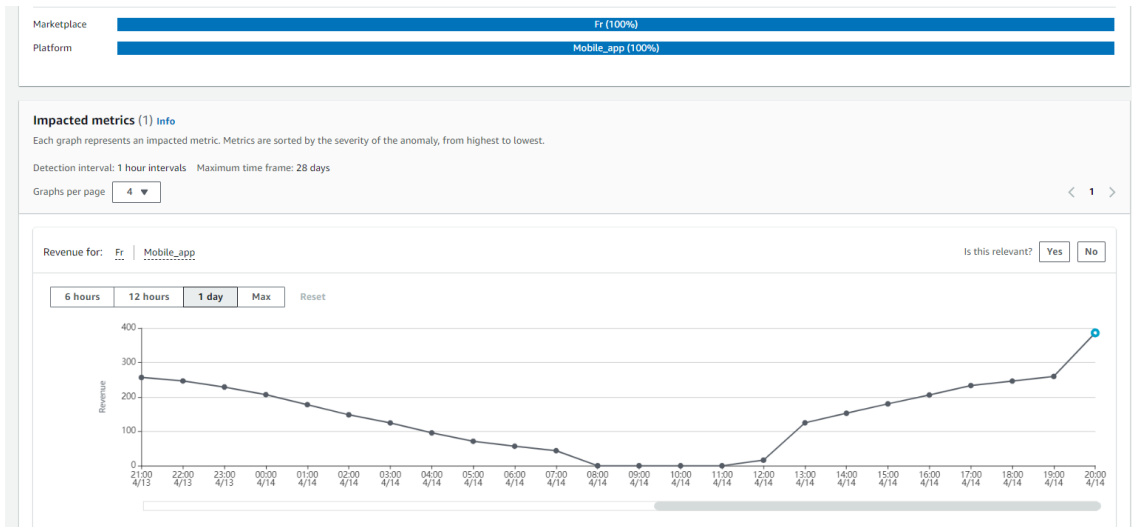
See which metrics are impacted, identify their root cause, and tune the results to improve accuracy.

View anomalies

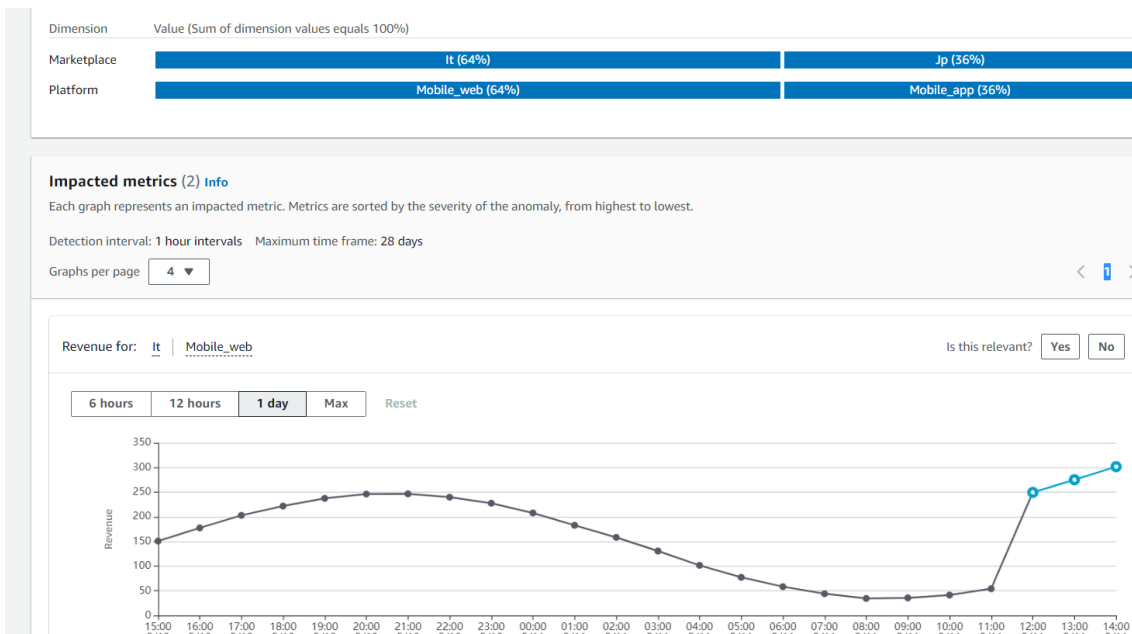
- [Severity threshold]のつまみを調整すると、異常値をもとに絞り込むことができます



- 異常は、中央値からの乖離であり、良い物悪いもの両方が含まれています。例えば以下の場合、なにがしかの理由でフランス売り上げが伸びたことがわかります。



多くの異常は単一 Dimensions により構成されていますが、中には以下のように複数要因が絡んでいるものも存在しています。



これにより更なる調査を行うための絞り込みなどが容易になります。

おつかれさまでした！削除は以下を行ってください

- Detector
- S3 バケット
- IAM Role

