
予測モデルを作成するスクリプト

リリース **1.0.0**

SIP-MI

2020 年 04 月 06 日

目次:

第 1 章	予測モデル値設定ツールの使い方	1
1.1	概要	1
1.2	はじめに	1
1.3	準備	1
1.4	使い方	1
1.4.1	引数、オプション	2
1.4.2	予測モデル値の作成スクリプトの使い方	2
1.5	まとめ	2
1.5.1	課題	2

第 1 章

予測モデル値設定ツールの使い方

1.1 概要

予測モデルの値が設定されていないワークフローに対して、入力ポートと出力ポートのデータ、ワークフロー名、説明文のデータを収集し、特性空間インベントリに予測モデルを定義するツールの使い方について記す。

1.2 はじめに

GPDB から情報を取得する際に、ワークフローに対して予測モデルの値が設定されていると、スムーズに情報が取り出せるとのこと。

予測モデルの値は、特性空間インベントリに予測モデルとして登録することで、予測モデルの値が発給される。

特性空間インベントリに予測モデルを登録するには、予測モデル名、予測モデルの説明、言語、入出力ポートの情報すべてが必要になる。これらの設定を用意するにあたり、入出力ポート数が多数あるワークフローも多く、設定するのに時間を要する。

なので、既存のワークフローからワークフロー名、ワークフローの説明、入出力ポートの情報を取得し、予測モデル名にワークフロー名、予測モデルの説明にワークフローの説明、入出力ポート一覧を自動で作成する python3.6 スクリプトを作成した。

作成した予測モデル値の作成スクリプトの説明を記す。

1.3 準備

CentOS 7.4 上に、python 3.6 の実行環境が必要。python のコマンドを python3.6 に切り替える必要はない。

1.4 使い方

コマンドラインインターフェースで実行する。

以下のコマンドを実行すると、

```
# python3.6 20200217_ut_config_prediction_model.py
```

MInt システムに対し、ユーザーのトークン情報を使って、サンプルワークフロー情報を特性空間インベントリに予測モデルとして登録する設定が実行される。

アクセス先やワークフローの情報が異なるので、次節に示す引数を指定して、動作させる対象を指定すること。

1.4.1 引数、オプション

以下に、スクリプトに与えられる引数を示す。

- -t、-token トークンを指定したものにする。
- -o、-host アクセス先の MI システムホストの指定する。
- -p、-port アクセス先 MI システムのポートを指定する (整数値のみ)。
- -v、-wfpiver WF-API のバージョン番号を指定する。
- -n、-wfpiname WF-API の名称が変わったら、変更後の名称を設定する。
- -u、-invapiver インベントリ API のバージョン番号を指定する。
- -m、-invapiname インベントリ API の名称が変わったら、変更後の名称を設定する。
- -l、-invupapiname インベントリ更新 API の名称が変わったら、変更後の名称を設定する。バージョンは、invapiver の値を引き継ぐ。
- -w、-wfid ワークフロー ID を設定する

1.4.2 予測モデル値の作成スクリプトの使い方

予測モデルの値を設定するスクリプトの変数定義を直接書き換えて使う方法と、引数を使って、定義を上書きする 2 種類の使い方がある。

変数を直接書き換える場合は、546 行目付近から始まる `## variables` 以降、560 行目付近の `## CONSTS` までは、右に有るコメントに沿って書き換える。この方法は、Python プログラミングに非常に詳しい方のみ利用すること。

スクリプト実行引数を与えて、動作を変更する。前節の引数説明に沿って、ホスト名やワークフローの ID を変更する。

```
# python3.6 20200217_ut_config_prediction_model.py -o <ホスト名> -p <ポート番号 (整数)> -t <トークン> -w <ワークフロー ID>
```

1.5 まとめ

プログラム引数で、アクセス先の URI やワークフロー名を指定することで、ワークフローの名称、説明、入出力ポートの情報をベースにインベントリに予測モデル情報を作成する、予測モデル値の作成スクリプトの説明を記した。

1.5.1 課題

- ワークフロー API でワークフローの情報が更新できない

ワークフロー API v2 には、ワークフローのアップデート機能が無く、ワークフローの追加しかできない同一機能のワークフローが新規ワークフローの爆発的な増加しても仕方がないので、ワークフローへの予測モデルの値指定は手動となっている。