

機械学習プラットフォーム・サービス調査  
報告書

平成31年4月

しまねソフト研究開発センター



# 目次

|              |   |           |
|--------------|---|-----------|
| <b>第 1 章</b> | <b>調査の目的</b>                                  | <b>5</b>  |
| <b>第 2 章</b> | <b>プラットフォームリスト (プラットフォーム名前順)</b>              | <b>6</b>  |
| 2.1          | Amazon / AWS(Amazon Web Service) . . . . .    | 6         |
| 2.1.1        | Amazon SageMaker . . . . .                    | 6         |
| 2.2          | Google / GCP(Google Cloud Platform) . . . . . | 6         |
| 2.2.1        | Cloud AutoML . . . . .                        | 6         |
| 2.2.2        | CLOUD MACHINE LEARNING ENGINE . . . . .       | 7         |
| 2.3          | IBM . . . . .                                 | 7         |
| 2.3.1        | Watson Machine Learning . . . . .             | 7         |
| 2.3.2        | Watson Studio . . . . .                       | 7         |
| 2.4          | Microsoft / Azure Cloud Services . . . . .    | 7         |
| 2.4.1        | Azure Machine Learning Studio . . . . .       | 7         |
| 2.4.2        | Azure Machine Learning サービス . . . . .         | 8         |
| 2.5          | Sony . . . . .                                | 8         |
| 2.5.1        | Neural Network Console . . . . .              | 8         |
| 2.6          | Uber . . . . .                                | 8         |
| 2.6.1        | Ludwig . . . . .                              | 8         |
| 2.7          | 今回調査できなかった機械学習プラットフォーム及びサービス . . . . .        | 9         |
| <b>第 3 章</b> | <b>利用したデータ</b>                                | <b>10</b> |
| 3.1          | 国土交通省 土地総合情報システム . . . . .                    | 10        |
| 3.1.1        | 土地価格データ (テーブルデータ) . . . . .                   | 10        |
| 3.1.2        | 土地価格データ内容 . . . . .                           | 10        |
| 3.2          | 中国電力でんき予報 . . . . .                           | 12        |
| 3.2.1        | 電力使用実績データ (時系列データ) . . . . .                  | 12        |

|              |   |           |
|--------------|---|-----------|
| 3.3          | Food-101 . . . . .                              | 13        |
| 3.3.1        | 料理画像データ (画像データ) . . . . .                       | 13        |
| <b>第 4 章</b> | <b>Amazon / AWS(Amazon Web Service)</b>         | <b>14</b> |
| 4.1          | Amazon SageMaker . . . . .                      | 14        |
| 4.1.1        | Amazon SageMaker 実際の使用感 . . . . .               | 14        |
| 4.1.2        | 公式チュートリアル . . . . .                             | 15        |
| 4.1.3        | 実行コード . . . . .                                 | 15        |
| <b>第 5 章</b> | <b>Google / GCP(Google Cloud Platform)</b>      | <b>16</b> |
| 5.1          | Cloud AutoML Vision . . . . .                   | 16        |
| 5.1.1        | Cloud AutoML Vision 実際の使用感 . . . . .            | 16        |
| 5.2          | CLOUD MACHINE LEARNING ENGINE . . . . .         | 17        |
| 5.2.1        | CLOUD MACHINE LEARNING ENGINE 実際の使用感 . . . . .  | 17        |
| 5.2.2        | 公式チュートリアル . . . . .                             | 18        |
| 5.2.3        | 実行コード . . . . .                                 | 18        |
| <b>第 6 章</b> | <b>IBM</b>                                      | <b>19</b> |
| 6.1          | Watson Machine Learning . . . . .               | 19        |
| 6.1.1        | Watson Machine Learning 実際の使用感 . . . . .        | 19        |
| 6.2          | Watson Studio . . . . .                         | 20        |
| 6.2.1        | Watson Studio 実際の使用感 . . . . .                  | 20        |
| 6.2.2        | 実行の流れ . . . . .                                 | 21        |
| <b>第 7 章</b> | <b>Microsoft / Azure Cloud Services</b>         | <b>22</b> |
| 7.1          | Azure Machine Learning Studio . . . . .         | 22        |
| 7.1.1        | Azure Machine Learning Studio 実際の使用感 . . . . .  | 22        |
| 7.1.2        | 公式チュートリアル . . . . .                             | 23        |
| 7.1.3        | 実行の流れ . . . . .                                 | 23        |
| 7.2          | Azure Machine Learning Service . . . . .        | 23        |
| 7.2.1        | Azure Machine Learning Service 実際の使用感 . . . . . | 23        |
| 7.2.2        | 公式チュートリアル . . . . .                             | 24        |
| 7.2.3        | 実行コード . . . . .                                 | 24        |
| <b>第 8 章</b> | <b>Sony</b>                                     | <b>25</b> |
| 8.1          | Neural Network Console . . . . .                | 25        |

---

|              |  |           |
|--------------|--|-----------|
| 8.1.1        | Neural Network Console 実際の使用感 . . . . .                            | 25        |
| 8.2          | 公式チュートリアル . . . . .  | 26        |
| <b>第 9 章</b> | <b>Uber</b>  | <b>28</b> |
| 9.1          | Ludwig . . . . .   | 28        |
| 9.1.1        | Ludwig 実際の使用感 . . . . .  | 28        |
| 9.1.2        | 公式チュートリアル . . . . .  | 28        |
| 9.1.3        | 実行コード . . . . .  | 29        |
| <b>付録 A</b>  | <b>未調査プラットフォームサービス</b>   | <b>30</b> |
| A.1          | DataRobot . . . . .  | 30        |
| A.1.1        | DataRobot の主な特徴 . . . . .  | 30        |
| A.1.2        | 調査対象から外した理由 . . . . .  | 30        |
| A.2          | Alibaba Cloud / Platform of Artificial Intelligence(PAI) . . . . . | 31        |
| A.2.1        | Alibaba Cloud PAI の主な特徴 . . . . .                                  | 31        |
| A.2.2        | 調査対象から外した理由 . . . . .  | 31        |
|              | <b>著者紹介</b>  | <b>32</b> |



# 第1章

## 調査の目的

近年は機械学習のプラットフォームやサービスも増え、類似している点も多く見受けられる。新しく機械学習プロジェクトを開始する際には、技術検討にも多くのコストがかかってしまう。そのため、今回の調査では実際に各サービスを触り使用感や実際のコードを記述することで、新規機械学習プロジェクトの技術選択の際に、どのプラットフォームやサービスが適しているのかを判断するための参考資料を作成した。

## 第 2 章

# プラットフォームリスト (プラットフォーム名前順)

今回は次のプラットフォーム及びサービスについて実際に利用して調査した。それぞれサービスの特徴を記述する。なお、2019 年 2 月頃に利用できる機能・環境での評価を行っている。

## 2.1 Amazon / AWS(Amazon Web Service)

### 2.1.1 Amazon SageMaker

- AWS のクラウド上に開発環境を作り Jupyter Notebook を作って作業ができる
- 学習用データや Jupyter Notebook の情報も S3 上に保管される
- 自動的に複数の機械学習アルゴリズムの構成を最適化してくれる
- 基本ニューラルネットワークのみ

## 2.2 Google / GCP(Google Cloud Platform)

### 2.2.1 Cloud AutoML

- GUI 上で画像分類のラベリング・モデル作成・デプロイまで行うことができる
- コーディングの必要が無い
- 既に学習済みのモデルを使えるため、比較的少ないデータ数でも画像分類ができる
- GCP のストレージバケット (Google Cloud Storage) へ画像をアップロードをおこなう

- Cloud AutoML には他にも Translation や Natural Language 等がある

## **2.2.2 CLOUD MACHINE LEARNING ENGINE**

- TensorFlow、XGBoost や Scikit-learn といったライブラリで学習できるコードを GCP のインスタンス上で高速に処理できる
- データ等は GCP のストレージバケット (Google Cloud Storage) へアップロードしておく
- 基本的にコンソール上で作業を行う

## **2.3 IBM**

### **2.3.1 Watson Machine Learning**

- Watson プラットフォームへデータをアップロードし、画面上での項目を設定するだけでモデルを構築できる
- コーディングの必要が無い

### **2.3.2 Watson Studio**

- GUI 上でブロックを組み立てることでデータの処理を記述できる
- コーディングの必要が無い
- Modler Flow と Neural Network Modeler が存在する
  - Modler Flow ではランダムフォレスト等の処理のフローを指定できる
  - Neural Network Modler ではニューラルネットワークの層を指定できる
    - \* 現在は自然言語処理または画像分類が可能

## **2.4 Microsoft / Azure Cloud Services**

### **2.4.1 Azure Machine Learning Studio**

- GUI 上でブロックを組み立てることでデータの処理を記述できる
- 都度処理実行をすることで、現在のデータの状態を確認しながら進めることができる

- コーディングの必要が無い

## 2.4.2 Azure Machine Learning サービス

- Microsoft Azure Notebooks というクラウド上で Jupyter Notebook を立ち上げて作業を行う
- データの処理は azureml-dataprep という独自の仕組みで Dataflow 型を利用する
- モデル構築もランダムフォレスト、SGD や LightGBM といった複数の分類器を利用して最終的にスタッキング処理を行う

## 2.5 Sony

### 2.5.1 Neural Network Console

- Windows 版とクラウド版が利用できる
- コーディングの必要が無い
- Windows 版はソフトウェアとして公開されており無料で利用できる
- クラウド版は従量課金制
- ローカルの PC で GPU が利用できる環境であれば容易に利用可能
- 画像データのトリミング処理等の機能もソフトウェアに付随する
- サンプルのプロジェクト等も画像分類処理が多い

## 2.6 Uber

### 2.6.1 Ludwig

- Python ライブラリのため無料で利用することができる
- コーディングの必要が無い
- データの説明変数と目的変数の情報を設定するだけで学習モデルを構築してくれる
- コンソール上で実行可能
- 画像の分類・ラベリング学習もできる

## 2.7 今回調査できなかった機械学習プラットフォーム及びサービス

次のサービスは今回の調査対象から外した。サービスの概要や今回調査対象とならなかった理由については付録にて記述する。

- DataRobot
- Alibaba Cloud / Platform of Artificial Intelligence(PAI)

## 第 3 章

# 利用したデータ

今回の調査では次のデータを利用して機械学習の処理を行った。

### 3.1 国土交通省 土地総合情報システム

ホームページ: <http://www.land.mlit.go.jp/webland/servlet/MainServlet>

#### 3.1.1 土地価格データ (テーブルデータ)

- 島根平成 29 年第 1 四半期～平成 30 年第 3 四半期

#### 3.1.2 土地価格データ内容

CSV ファイルで次の項目が存在する

- No
  - 通し番号
- 種類
  - 不動産の種類
  - 宅地 (土地)、宅地 (土地と建物)、中古マンション等、農地、林地
- 地域
  - 宅地 (土地)、宅地 (土地と建物) の時に、周辺地域の状況
  - 住宅地、商業地、工業地、宅地見込地
- 市区町村コード
  - 一意の数字 (例: '32201')
- 都道府県名

- 市区町村名
- 地区名
- 最寄駅：名称
  - － 宅地（土地）、宅地（土地と建物）、中古マンション等の時に、最寄りの鉄道駅
- 最寄駅：距離（分）
  - － 30分未満は分単位、30分以上は30分以上1時間未満、1時間以上1時間30分未満、1時間30分以上2時間未満、2時間以上に区分して表示
- 取引価格（総額）
  - － 宅地（土地）、宅地（土地と建物）、中古マンション等、農地、林地の時の取引の総額（万円）
- 坪単価
  - － 宅地（土地）の時に、土地の取引総額を土地面積（坪）で除した坪当たりの価格（万円）を表示
- 間取り
  - － 中古マンション等の時に、専有部分の間取りを表示
  - － 例) 1LDK、3LDK
- 面積（㎡）
  - － 宅地（土地）、宅地（土地と建物）、農地、林地の時に、取引当事者への調査により得られた土地の実測面積（㎡）を、実測面積が不詳の場合には登記簿に記載されている公簿面積（㎡）を表示
- 取引価格（㎡単価）
  - － 宅地（土地）の時に、土地の取引総額を土地面積（㎡）で除した㎡当たりの価格（万円）を表示
- 土地の形状
  - － 正方形、ほぼ正方形、長方形、ほぼ長方形、台形、ほぼ台形、不整形、ほぼ整形、袋地等
- 間口
- 延床面積（㎡）
- 建築年
- 建物の構造
  - － SRC、RC、鉄骨造、軽量鉄骨造、ブロック造、木造
- 用途
  - － 例) 住宅、事務所、店舗、工場、倉庫、作業場、駐車場
- 今後の利用目的
  - － 例) 住宅、店舗、事務所、工場、倉庫

- 前面道路：方位
- 前面道路：種類
  - － 例) 国道、都道府県道、市町村道、区画街路、私道
- 前面道路：幅員 (m)
- 都市計画
  - － 1 低住専、2 低住専、1 中住専、2 中住専、1 種住居、2 種住居、準住居、田園住居、近隣商業、商業、準工業、工業、工業専用、調整区域、非線引き、準都計、都計外
- 建ぺい率 (%)
- 容積率 (%)
- 取引時点
  - － 取引当事者への調査により得られた売買契約日を、売買契約日が不詳の場合には登記簿に記載されている登記原因日を四半期 (3ヶ月) 単位で表示
- 改装
  - － 改装したものは「改装済」、改装していないものは「未改装」
- 取引の事情等
  - － 取引にあたって価格に影響があると思われる追加情報を表示

### 3.2 中国電力でんき予報

ホームページ: <http://www.energia.co.jp/jukyuu/>

#### 3.2.1 電力使用実績データ (時系列データ)

- 2017年4月1日～2019年2月15日分まで時間毎の電力使用実績

▼リスト 3.1 juyo-2017-2018.csv

```
DATE, TIME, 実績 (万 kW)
2017/4/1, 0:00, 654
2017/4/1, 1:00, 660
2017/4/1, 2:00, 685
...
```

## 3.3 Food-101

ホームページ: [https://www.vision.ee.ethz.ch/datasets\\_extra/food-101/](https://www.vision.ee.ethz.ch/datasets_extra/food-101/)

### 3.3.1 料理画像データ (画像データ)

- 101 の料理カテゴリで合計 101000 枚の画像用意されている
  - 料理カテゴリ例) apple\_pie, gyoza, pizza, sushi, waffles

## 第 4 章

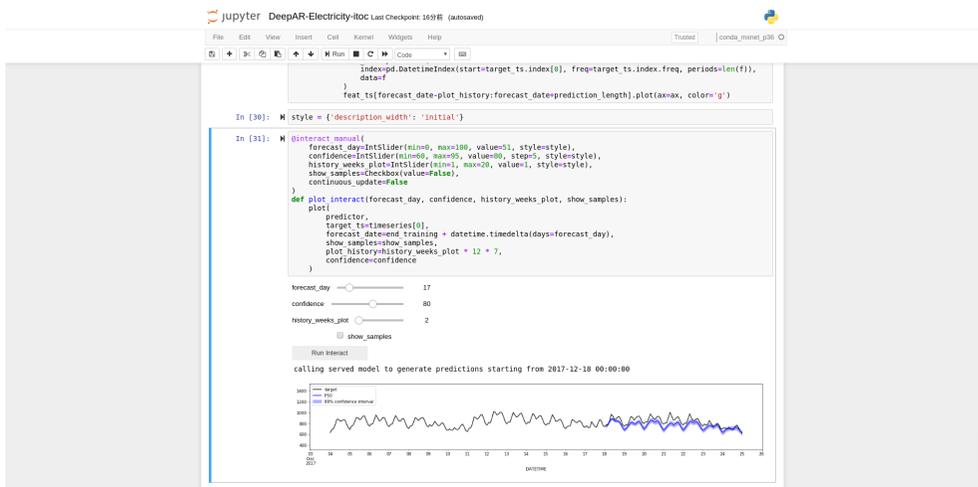
# Amazon / AWS(Amazon Web Service)

### 4.1 Amazon SageMaker

公式ページ: <https://aws.amazon.com/jp/sagemaker/>

#### 4.1.1 Amazon SageMaker 実際の使用感

AWS で Jupyter Notebook インスタンスを立ち上げてモデル作成を行い、そのまま学習したモデルをデプロイすることで予測をできるようにする。AWS 上で S3 のデータバケットを作成し、そこにデータをアップロードし、Jupyter Notebook 等のデータもそこに保存される。検証では電力使用実績データで予測を行った。DeepAR という SageMaker で用意されているものを利用する。



▲ 図 4.1 Amazon SageMaker: Jupyter Notebook での予測結果

## 4.1.2 公式チュートリアル

[https://docs.aws.amazon.com/ja\\_jp/sagemaker/latest/dg/algos.html](https://docs.aws.amazon.com/ja_jp/sagemaker/latest/dg/algos.html)

## 4.1.3 実行コード

<https://github.com/regomm/itoc-codes/blob/master/aws-sagemaker.ipynb>

## 第 5 章

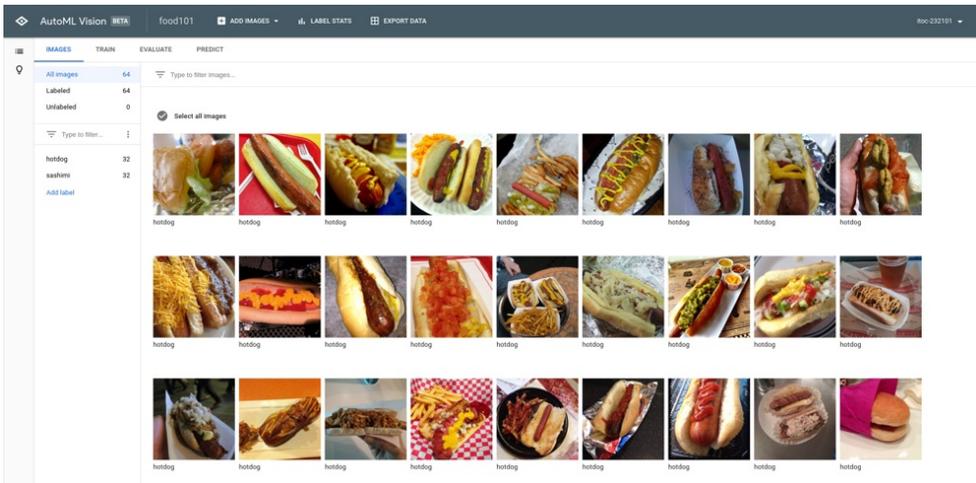
# Google / GCP(Google Cloud Platform)

### 5.1 Cloud AutoML Vision

公式ページ: <https://cloud.google.com/automl/?hl=ja>

#### 5.1.1 Cloud AutoML Vision 実際の使用感

全てが GUI 上で作業ができ、とても簡単にモデルを構築して Web API として利用できるのと、データが少なくても (1 つのラベルに対して 10 枚程度でも) 画像分類を行うことができる。一つの画像に対して複数ラベルの予測も対応しているが、その場合はより多くの画像が必要になる。全ての画像に GUI 上でラベルを設定する必要があるが、1~5 日間の所要時間で Google ヒューマンラベリングサービスを利用することで人の手によるラベル付けを行うことも可能。料理画像データを登録して予測を行った。



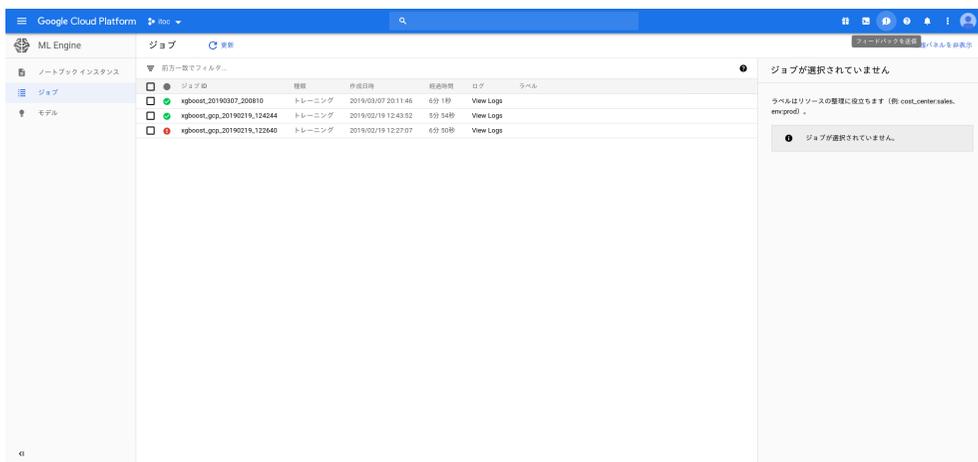
▲図 5.1 Cloud AutoML Vision: ラベル設定画面

## 5.2 CLOUD MACHINE LEARNING ENGINE

公式ページ: <https://cloud.google.com/ml-engine/?hl=ja>

### 5.2.1 CLOUD MACHINE LEARNING ENGINE 実際の使用感

既に XGBoost 等のライブラリで学習できるコードがあれば、それを GCP のクラウド上でモデル作成を行うことができ、そのままモデルのデプロイも可能になる。全ての操作をコマンドライン上で行う。先にローカルで実行を確認してからクラウドへデプロイすることができる。土地価格データを XGBoost を用いて、モデル作成と予測を行った。



▲図 5.2 CLOUD MACHINE LEARNING ENGINE: クラウドで実行したジョブ一覧画面

### 5.2.2 公式チュートリアル

<https://cloud.google.com/ml-engine/docs/scikit/getting-started-training?hl=ja>

### 5.2.3 実行コード

<https://github.com/regonn/itoc-codes/blob/master/gcp.ipynb>

# 第 6 章

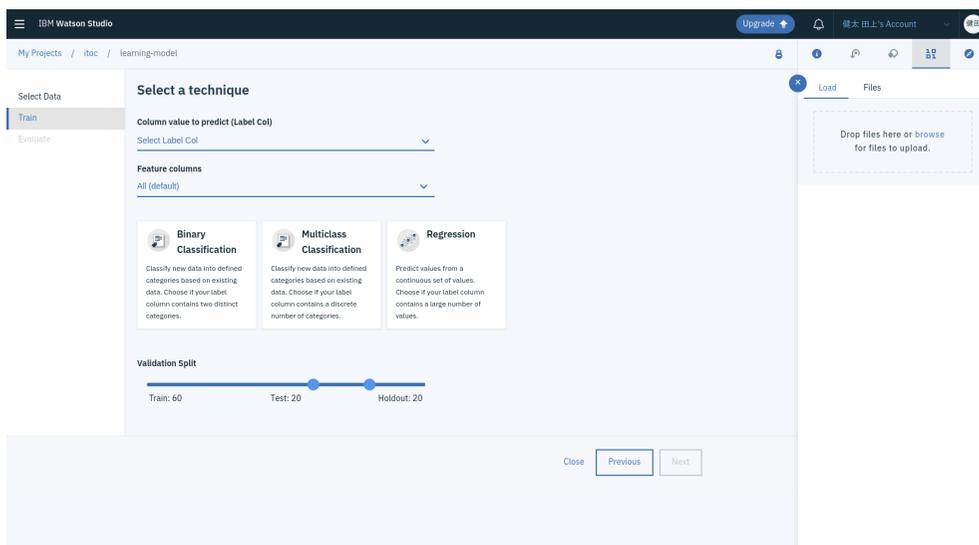
# IBM

## 6.1 Watson Machine Learning

公式ページ: <https://www.ibm.com/jp-ja/cloud/machine-learning>

### 6.1.1 Watson Machine Learning 実際の使用感

プラットフォームヘデータのアップロードを行い、予測する項目と、予測の種類を選択するだけで学習ができ、モデルをデプロイすることができる。



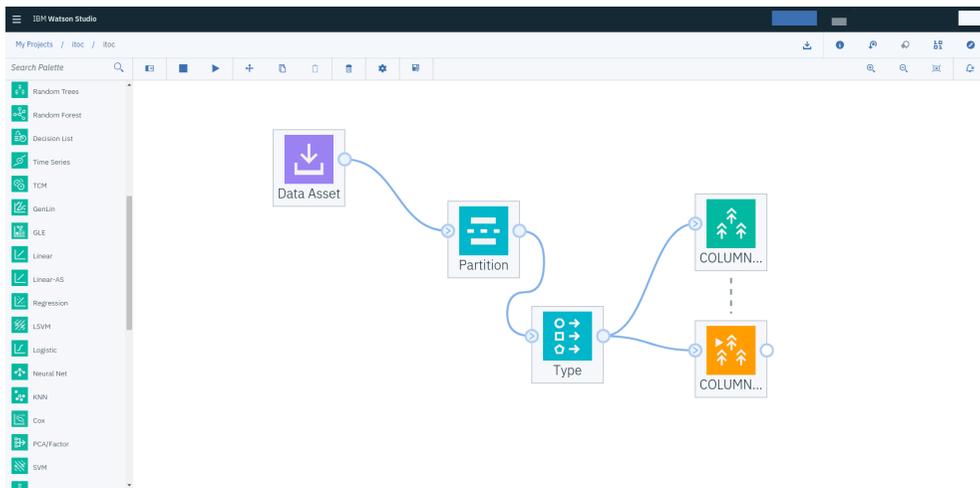
▲図 6.1 Watson Machine Learning: モデル設定画面

## 6.2 Watson Studio

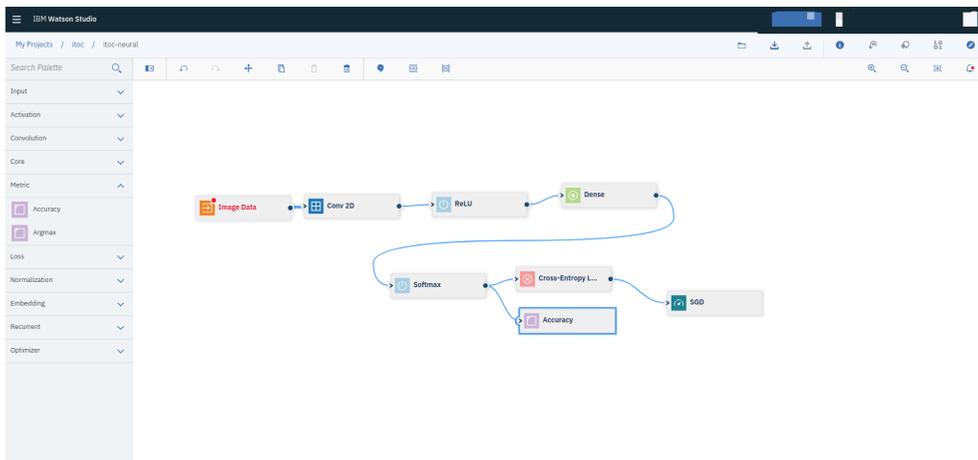
公式ページ: <https://www.ibm.com/jp-ja/cloud/machine-learning>

### 6.2.1 Watson Studio 実際の使用感

ブロックを組み立てるだけでモデル作成まで行うことができる。同じ仕組みで、機械学習のフローや深層学習レイヤーを設定することができる。



▲図 6.2 Watson Studio: Modler Flow の設定画面



▲ 図 6.3 Watson Studio: Neural Network Modeler の設定画面

### 6.2.2 実行の流れ

<https://github.com/regonn/itoc-codes/blob/master/watson-studio.ipynb>

## 第7章

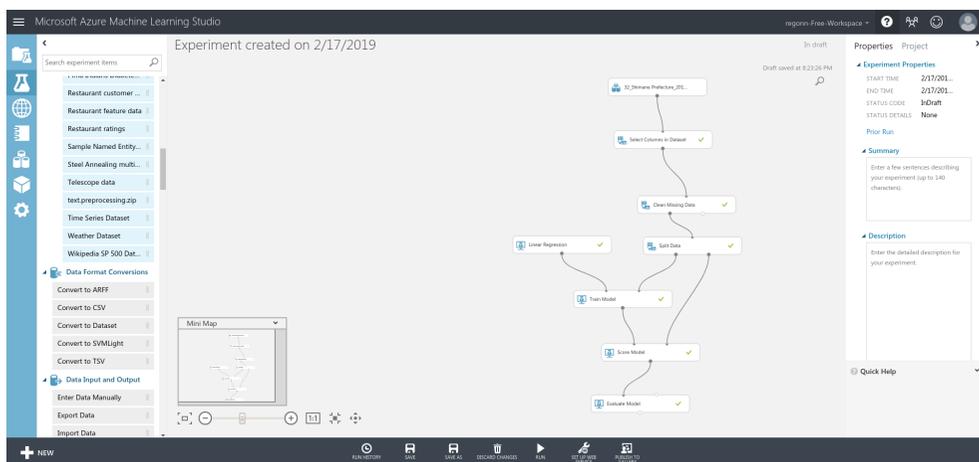
# Microsoft / Azure Cloud Services

## 7.1 Azure Machine Learning Studio

公式ページ: <https://azure.microsoft.com/ja-jp/services/machine-learning-studio/>

### 7.1.1 Azure Machine Learning Studio 実際の使用感

ブロックを組み立ててロジックを構築していく。現在のブロックへどういったデータが入ってくるのかも、ブロックを選択することで確認しながら構築していくことができる。



▲ 図 7.1 Azure Machine Learning Studio: モデル構築画面

### 7.1.2 公式チュートリアル

<https://docs.microsoft.com/ja-jp/azure/machine-learning/studio/create-experiment>

### 7.1.3 実行の流れ

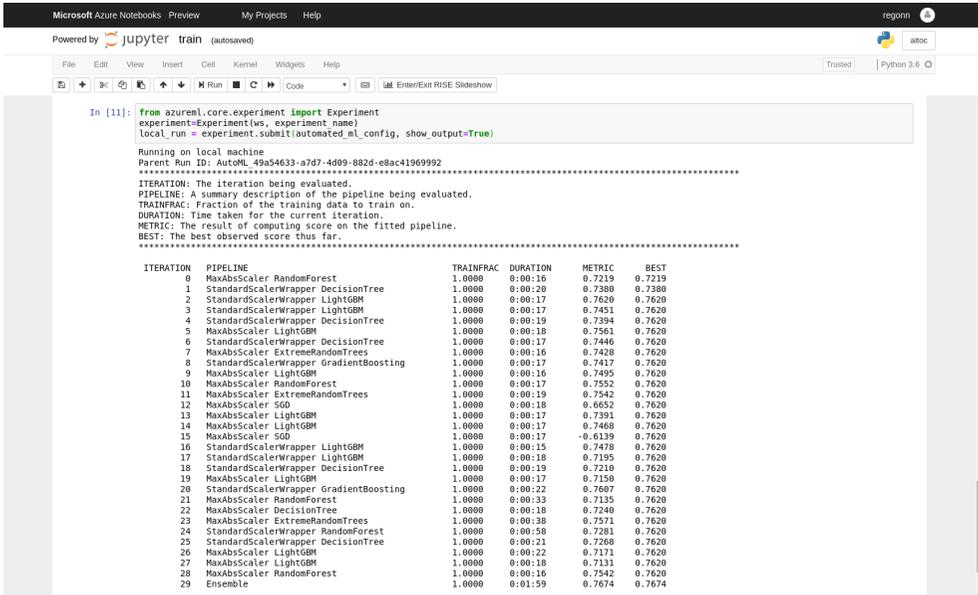
<https://github.com/regonn/itoc-codes/blob/master/azure-studio.ipynb>

## 7.2 Azure Machine Learning Service

公式ページ: <https://azure.microsoft.com/ja-jp/services/machine-learning-service/>

### 7.2.1 Azure Machine Learning Service 実際の使用感

Azure 上で Jupyter Notebook 用のインスタンスを立ち上げて、そこでモデル作成を行いデプロイまでできる。独自のライブラリでデータ処理を行うため慣れが必要だが、学習部分は色々な分類器を用いて学習を行ってくれるため、データを渡せばお手軽にモデル構築も可能になる。土地価格データを用いて予測を行った。



```
In [11]: from azureml.core.experiment import Experiment
experiment=Experiment(ws, experiment name)
local_run = experiment.submit(automated_ml_config, show_output=True)

Running on local machine
Parent Run ID: AutoML_49x54633-a7d7-4d99-882d-e8ac41969992
*****
ITERATION: The iteration being evaluated.
PIPELINE: A summary description of the pipeline being evaluated.
TRAINFRAC: Fraction of the training data to train on.
DURATION: Time taken for the current iteration.
METRIC: The result of computing score on the fitted pipeline.
BEST: The best observed score thus far.
*****
ITERATION PIPELINE TRAINFRAC DURATION METRIC BEST
0 MaxAbsScaler RandomForest 1.0000 0:00:16 0.7219 0.7219
1 StandardScalerWrapper DecisionTree 1.0000 0:00:20 0.7380 0.7380
2 StandardScalerWrapper LightGBM 1.0000 0:00:17 0.7620 0.7620
3 StandardScalerWrapper LightGBM 1.0000 0:00:17 0.7451 0.7620
4 StandardScalerWrapper DecisionTree 1.0000 0:00:19 0.7394 0.7620
5 MaxAbsScaler LightGBM 1.0000 0:00:18 0.7561 0.7620
6 StandardScalerWrapper DecisionTree 1.0000 0:00:17 0.7446 0.7620
7 MaxAbsScaler ExtremeRandomTrees 1.0000 0:00:16 0.7428 0.7620
8 StandardScalerWrapper GradientBoosting 1.0000 0:00:17 0.7417 0.7620
9 MaxAbsScaler LightGBM 1.0000 0:00:16 0.7495 0.7620
10 MaxAbsScaler RandomForest 1.0000 0:00:17 0.7552 0.7620
11 MaxAbsScaler ExtremeRandomTrees 1.0000 0:00:19 0.7542 0.7620
12 MaxAbsScaler LightGBM 1.0000 0:00:18 0.6652 0.7620
13 MaxAbsScaler SGD 1.0000 0:00:17 0.7391 0.7620
14 MaxAbsScaler LightGBM 1.0000 0:00:17 0.7468 0.7620
15 MaxAbsScaler SGD 1.0000 0:00:17 0.6339 0.7620
16 StandardScalerWrapper LightGBM 1.0000 0:00:15 0.7478 0.7620
17 StandardScalerWrapper LightGBM 1.0000 0:00:18 0.7195 0.7620
18 StandardScalerWrapper DecisionTree 1.0000 0:00:19 0.7210 0.7620
19 MaxAbsScaler LightGBM 1.0000 0:00:17 0.7150 0.7620
20 StandardScalerWrapper GradientBoosting 1.0000 0:00:22 0.7607 0.7620
21 MaxAbsScaler RandomForest 1.0000 0:00:13 0.7155 0.7620
22 MaxAbsScaler DecisionTree 1.0000 0:00:18 0.7240 0.7620
23 MaxAbsScaler ExtremeRandomTrees 1.0000 0:00:18 0.7571 0.7620
24 StandardScalerWrapper RandomForest 1.0000 0:00:18 0.7281 0.7620
25 StandardScalerWrapper DecisionTree 1.0000 0:00:21 0.7268 0.7620
26 MaxAbsScaler LightGBM 1.0000 0:00:22 0.7171 0.7620
27 MaxAbsScaler LightGBM 1.0000 0:00:18 0.7131 0.7620
28 MaxAbsScaler RandomForest 1.0000 0:00:16 0.7542 0.7620
29 Ensemble 1.0000 0:01:59 0.7674 0.7674
```

▲ 図 7.2 Azure Machine Learning Service: 学習時画面

### 7.2.2 公式チュートリアル

<https://docs.microsoft.com/ja-jp/azure/machine-learning/service/tutorial-data-prep>

### 7.2.3 実行コード

<https://github.com/regonn/itoc-codes/blob/master/azure.ipynb>

## 第 8 章

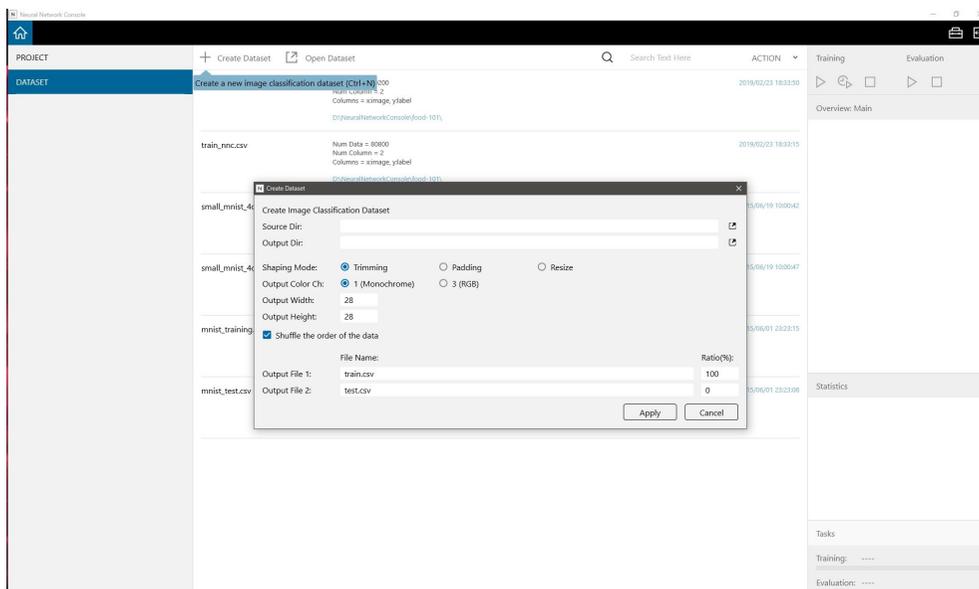
# Sony

### 8.1 Neural Network Console

公式ページ: <https://dl.sony.com/ja/>

#### 8.1.1 Neural Network Console 実際の使用感

無料で利用できるソフトウェアとして配布されており、ブロックを用いて、深層学習モデルのレイヤーを構築することができる。手元にある Windows のパソコンで実行できるため、GPU 等を利用しやすい。テーブルデータも利用できるが、Neural Network Console では画像系のサンプルプロジェクトが色々と用意されており、自分で画像データセットを作成する機能もついている。



▲ 図 8.1 Neural Network Console: データセット作成画面

## 8.2 公式チュートリアル

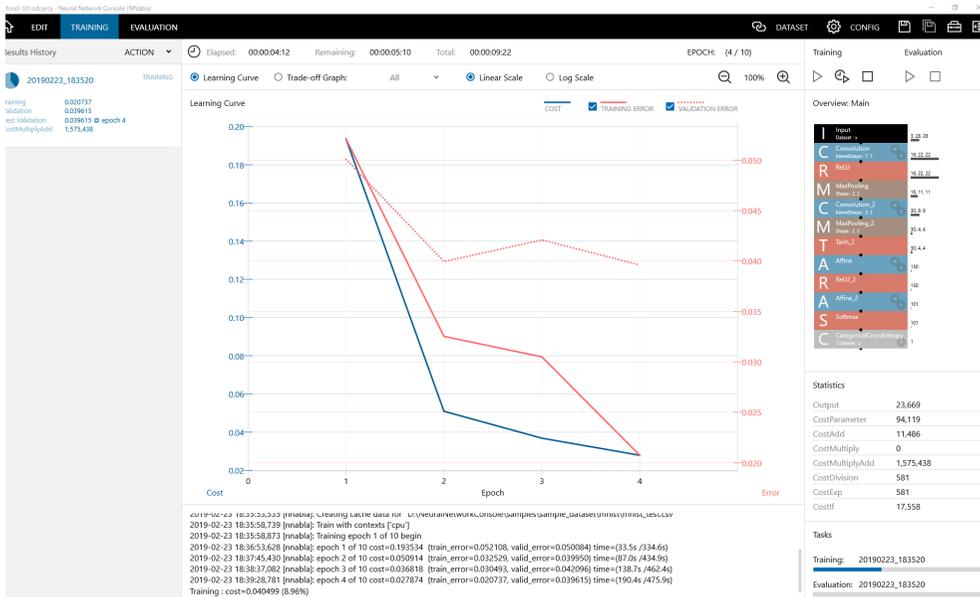
<https://support.dl.sony.com/docs-ja/>

画像サイズ処理も、トリミング、パディング、リサイズから選べたり出力したい画像サイズも指定できる。画像も Food-101 のようにラベル毎に分類されているディレクトリを選択するだけで良い。トレーニングデータとテストデータの比率等も選ぶことができる。データセット作成が終わると次のような csv が出力される。

### ▼リスト 8.1 train.csv

```
x:image,y:label
.\beignets\2128508.png,6
.\beet_salad\206201.png,5
.\grilled_cheese_sandwich\2103806.png,49
.\ceviche\1248987.png,15
.\spaghetti_carbonara\827576.png,91
...
```

ニューラルネットワークの層をブロックで組み合わせて学習を行う。



▲ 図 8.2 Neural Network Console: 学習時画面

## 第 9 章

# Uber

### 9.1 Ludwig

公式ページ: <https://uber.github.io/ludwig/>

#### 9.1.1 Ludwig 実際の使用感

無料で利用できる Python ライブラリとして配布されており、説明変数と目的変数の設定を記述するだけでモデル作成が可能。2019 年 2 月 11 日に最初のリリースが行われたため、まだ日本語の資料などは少ないが、公式のチュートリアルだけでも、何ができるかわかりやすい。Python コード上での処理やコンソールだけでもモデル作成をすることができる。

#### 9.1.2 公式チュートリアル

<https://uber.github.io/ludwig/examples/>

画像分類では次のような path とラベルの項目を持った csv ファイルを用意する。Food-101 の画像を縦横比が同じ大きさ (28 × 28px) になるように変換してある。

▼リスト 9.1 train.csv

```
image,label
converted_images/churros/1004234.png,churros
converted_images/churros/1013460.png,churros
...
```

そして、モデルを定義した model\_definition.yaml も用意する。

## ▼リスト 9.2 model\_definition.yaml

```
input_features:
- name: image
  type: image
  encoder: stacked_cnn

output_features:
- name: label
  type: category
```

実行はコマンドライン上で行う場合は次のようなコマンドを入力する。

```
$ ludwig experiment \
--data_csv train.csv \
--model_definition_file model_definition.yaml
```

これだけで、学習が開始されて、出来上がったモデルから予測も可能になる。  
今度は土地価格データを Python コード上で学習処理を行う。

### 9.1.3 実行コード

<https://github.com/regonn/itoc-codes/blob/master/ludwig.ipynb>

## 付録 A

# 未調査プラットフォームサービス

次のサービスについては今回の調査対象から外した。

### A.1 DataRobot

公式ページ: <https://www.datarobot.com/jp/>

#### A.1.1 DataRobot の主な特徴

- アップロードされたデータに対して特徴抽出やモデル作成まで行ってくれる
- 数多くの手法を並列処理させて高速に最も良い結果を出す手法を特定する

#### A.1.2 調査対象から外した理由

- DataRobot 社に問い合わせたところ、現在は試験的に触れる環境は存在しないとのこと
- そのため、実際に利用するには 1 年契約のプランを申し込む必要があるため調査対象から外した
- 実際に試してみたい場合には東京や大阪で定期的にワークショップが開催されている

## **A.2 Alibaba Cloud / Platform of Artificial Intelligence(PAI)**

### **A.2.1 Alibaba Cloud PAI の主な特徴**

- ブロックを組み立てることでデータの処理を記述できる

### **A.2.2 調査対象から外した理由**

- 現在日本サイトでは PAI はリリースされておらず、中国か国際アカウントでの利用が必要
- 日本語の資料も少なく、プロジェクトへの導入のしやすさという観点で調査対象から外した

# 著者紹介

田上 健太 / **Twitter: @regonn\_haizine**

島根県松江市で活動しているフリーランスエンジニア

Podcast regonn&curry.fm で最新の機械学習事情等を配信している



機械学習プラットフォーム・サービス調査報告書

2019年3月25日 初版 v1.0.0

著者 田上 健太

監修 しまねソフト研究開発センター 専門研究員 木村 忍  
" 専門研究員 高木 丈智

制作 しまねソフト研究開発センター

(C) 2019 ITOC