

NEC Advanced Analytics Platform V1.5 マネージドサービス サイジングガイド (簡易版)

RAPID 機械学習 時系列数値解析版編 (第1版)

本シートでは RAPID 機械学習の時系列数値解析版の分析に必要なメモリー、ストレージ容量の参考値を説明します。

時系列数値解析版には、1DCNN アルゴリズム、1DDCN アルゴリズム、1DOCN アルゴリズムによる分析の3種類があります。いずれのアルゴリズムも分析する入力データを行、列に分離して分析します。

表1には1DDCN アルゴリズムにおける6ケースの行、列、データサイズにおいて1回学習を行う場合のメモリー、ストレージ使用量の実測値を示します。

表2には同様に1DOCN アルゴリズムにおける1回学習を行う場合の実測値を示します。

表3、表4にはそれぞれ1回予測を行う場合のメモリー、ストレージ使用量の実測値を示します。

1DCNN アルゴリズムの実測値はありませんが、1DCNN を用いる場合のメモリー使用量とストレージ使用量は、一般に1DDCN アルゴリズムと同等かそれ以下となるため、1DDCN の場合を基準にメモリー使用量とストレージ使用量を見積もります。

これらの情報をもとにお客様が分析されようと考えている入力データの行、列、データサイズにと照らし合わせて必要なメモリー、ストレージ量を推測して見積もり、最適なプランを選択します。

表1 学習処理1回当たりの実測値(1DDCN アルゴリズム)

No	入力データ行数	入力データ列数	入力データサイズ	メモリー使用量	ストレージ使用量
1	10,000	10	0.26MB	0.30MB	4,312KB
2	10,000	50	1.05MB	0.46MB	4,348KB
3	10,000	100	2.06MB	0.50MB	4,392KB
4	100,000	10	2.69MB	0.37MB	4,312KB
5	100,000	50	10.69MB	0.67KB	4,348KB
6	100,000	100	20.69MB	0.77MB	4,352KB

表2 学習処理1回当たりの実測値(1DOCN アルゴリズム)

No	入力データ行数	入力データ列数	入力データサイズ	メモリー使用量	ストレージ使用量
1	10,000	10	0.26MB	0.13MB	1,244KB
2	10,000	50	1.05MB	0.56MB	1,304KB
3	10,000	100	2.06MB	1.01MB	1,396KB
4	100,000	10	2.69MB	0.81MB	1,244KB
5	100,000	50	10.69MB	2.67MB	1,304KB
6	100,000	100	20.69MB	9.58MB	1,396KB

表3 予測処理1回当たりの実測値(1DDCN アルゴリズム)

No	入力データ行数	入力データ列数	入力データサイズ	メモリー使用量	ストレージ使用量
1	10,000	10	0.26MB	0.04MB	1,404KB
2	10,000	50	1.05MB	0.08MB	1,596KB
3	10,000	100	2.06MB	0.15MB	1,792KB
4	100,000	10	2.69MB	0.14MB	19,502KB
5	100,000	50	10.69MB	0.44MB	18,092KB
6	100,000	100	20.69MB	1.05MB	18,096KB

表 4 予測処理 1 回当たりの実測値(1DOCN アルゴリズム)

No	入力データ行数	入力データ列数	入力データサイズ	メモリー使用量	ストレージ使用量
1	10,000	10	0.26MB	0.03MB	488KB
2	10,000	50	1.05MB	0.08MB	492KB
3	10,000	100	2.06MB	0.15MB	492KB
4	100,000	10	2.69MB	0.03MB	7,756KB
5	100,000	50	10.69MB	0.47MB	7,732KB
6	100,000	100	20.69MB	1.08MB	7,788KB

前提条件・注意事項等：

- 各アルゴリズムの測定条件は以下の通りです。
 - 1DDCN アルゴリズム
 - 学習モード：分類
 - 学習・予測コンフィグのパラメータ
 - WINDOW_SIZE: 256 (ウィンドウサイズ (一度に抜き出して入力するデータの行数))
 - TRAIN_EPOCH: 10 (学習の反復回数 (epoch 数))
 - その他のパラメータはデフォルト値
 - 1DOCN アルゴリズム
 - 学習モード：分類
 - 学習・予測コンフィグのパラメータ
 - WINDOW_SIZE: 256
 - 1DOCN_PREDICT_WINDOW_SIZE: 128 (判断区間のサイズ)
 - TRAIN_EPOCH: 10
 - その他のパラメータはデフォルト値
- 表 1、表 2 は、学習モードが「分類」の場合の学習処理の実測値ですが、学習モードが「回帰」の場合でもストレージ使用量、メモリー使用量とも「分類」とほぼ同等です。
- 表 3、表 4 は、学習モードが「分類」の場合の予測処理の実測値ですが、学習モードが「回帰」の場合、ストレージ使用量は「分類」の時より小さくなります。
- 学習・予測処理を行う前に前処理としてデータの加工を行う場合は、別途前処理済のデータを格納するストレージ容量も必要になります。
- 本書で示す実測値はあくまで目安としてご利用ください。お客様の分析データや分析方法によってはメモリー使用量やストレージ使用量が想定を上回る可能性があります。エントリープラン、エントリープラスプランでは後からメモリーを追加できないため、およびエントリープランでは後からストレージを追加できないため、不安な場合は 1 ランク上のプランを選択される等ご検討ください。より確からしい見積もりを出すためには、本番運用と同等の特性を持つサンプルデータを用いるなどして実測することを推奨します。

用語説明：

1DCNN	1 次元畳み込みニューラルネットワーク (1-Dimension Convolutional Neural Network) の略。 ニューラルネットワークの手法であり、画像解析に用いられる CNN を、時系列数値データ解析用にカスタマイズした手法。RAPID 機械学習(時系列数値解析版)で利用できるアルゴリズムの一つ。
1DDCN	1-Dimension Deep Convolutional Neural Network の略。 1D CNN よりもさらに深いネットワークでの高精度な学習を可能にする。RAPID 機械学習(時系列数値解析版)で利用できるアルゴリズムの一つ。
1DOCN	1-Dimension OneClass Neural Network の略。 正常データのみによる学習を可能にする。RAPID 機械学習(時系列数値解析版)で利用できるアルゴリズムの一つ。

本書について：

- より詳細な見積もりが必要な場合は、「NEC Advanced Analytics Platform V1.5 マネージドサービス サイジングガイド」を参照してください。
- 本書を事業者の許可なくコピーおよびその配布、ホームページへの掲載を禁じます。