

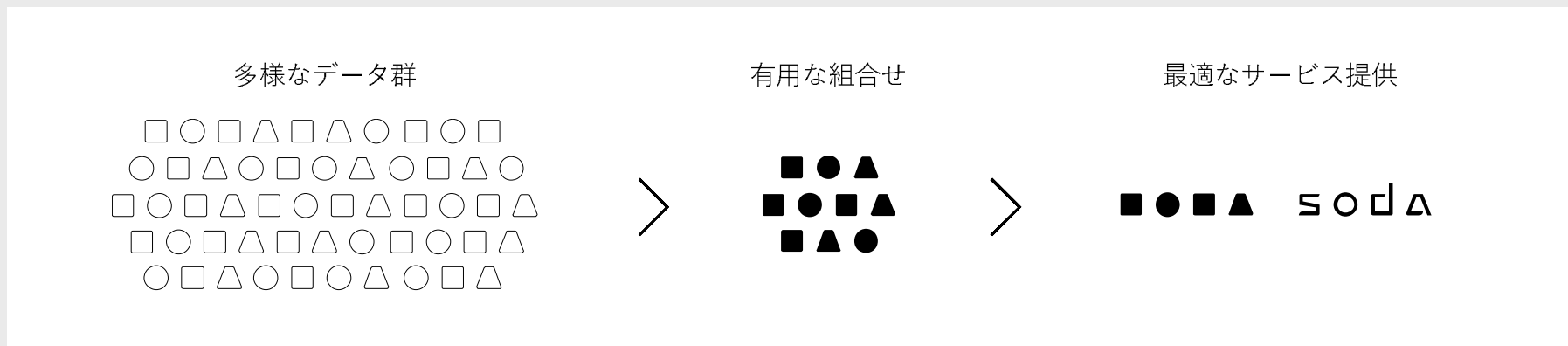
# VMS・TMS・Deep Learnerで サクッと自然言語処理

soda



## ビジネスのあらゆる局面においてデータ収集・利活用を促進

私たちは、クラウドやAIといった先端テクノロジーやアルゴリズム、ノウハウを活用し、マーケティング、セールス、オペレーション、バックオフィスなど、ビジネスのあらゆる局面においてデータ収集・利活用を促進。多様なデータの中から有用なものを組合せ、最適なサービスを提供します。



商号：株式会社soda（英語表記：soda Inc.）

所在地：〒461-0005 名古屋市東区東桜2-11-4 SHOEI SQUARE 4F

設立：令和元年7月1日

代表者：代表取締役社長 西川栄一

連絡先：TEL 052-880-7782 / FAX 052-880-7783

URL：nico-soda.jp

グループ会社：西川コミュニケーションズ株式会社

株式会社ディ・スタイル西川

Nishikawa Communications India Pvt. Ltd.

Nishikawa Communications Singapore Pte. Ltd.



## 簡易アンケート機能

フリーアンサーを活用できないか  
何か新しい示唆を得られないか



- ・ 件数：約4,600（FA回答者のみ）
- ・ 期間：約1年分
- ・ 内容：性別 / 年代 / 店舗名 /  
買い物の頻度 /  
ポイントカードの有無 /  
店舗へのご意見（FA）

※同一人物による複数回 回答が可能

## テキストマイニング手順【例】

[1] 分かち書き

[2] 辞書の整理・作成

手間が掛かる[2][4]の作業  
Deep Learning(AI)で  
何とかならないか!?

[3] 集計・分析・可視化

単語頻度解析 / 係り受け頻度解析 / 特徴語・特徴表現抽出 etc.

文章分類 / ことばネットワーク / グルーピング

評判抽出 (ポジネガ)

[4] 結果の読込み・まとめ

# テキストマイニング編



- [1] 分かち書き
- [2] 辞書の整理・作成
- [3] 集計・分析・可視化
- [4] 結果の読み込み

## オリジナルの辞書を作成します

ユーザ辞書

類義語辞書

分割辞書

分析の精度を上げるためには必須の作業  
ですが手間がかかります

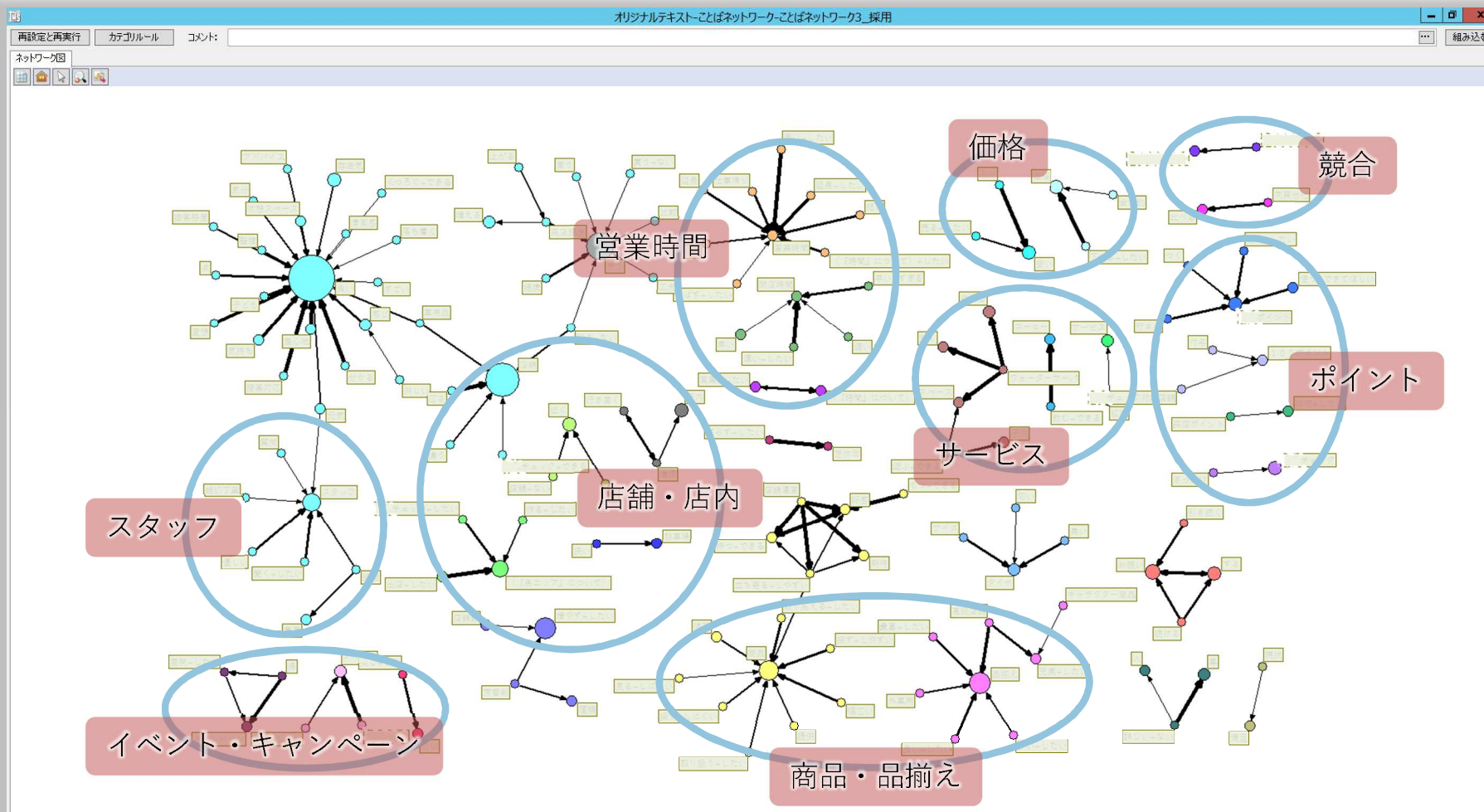
# テーマ（特徴）の考察



ことばネットワークなどを使い  
読み込んでテーマを設定します

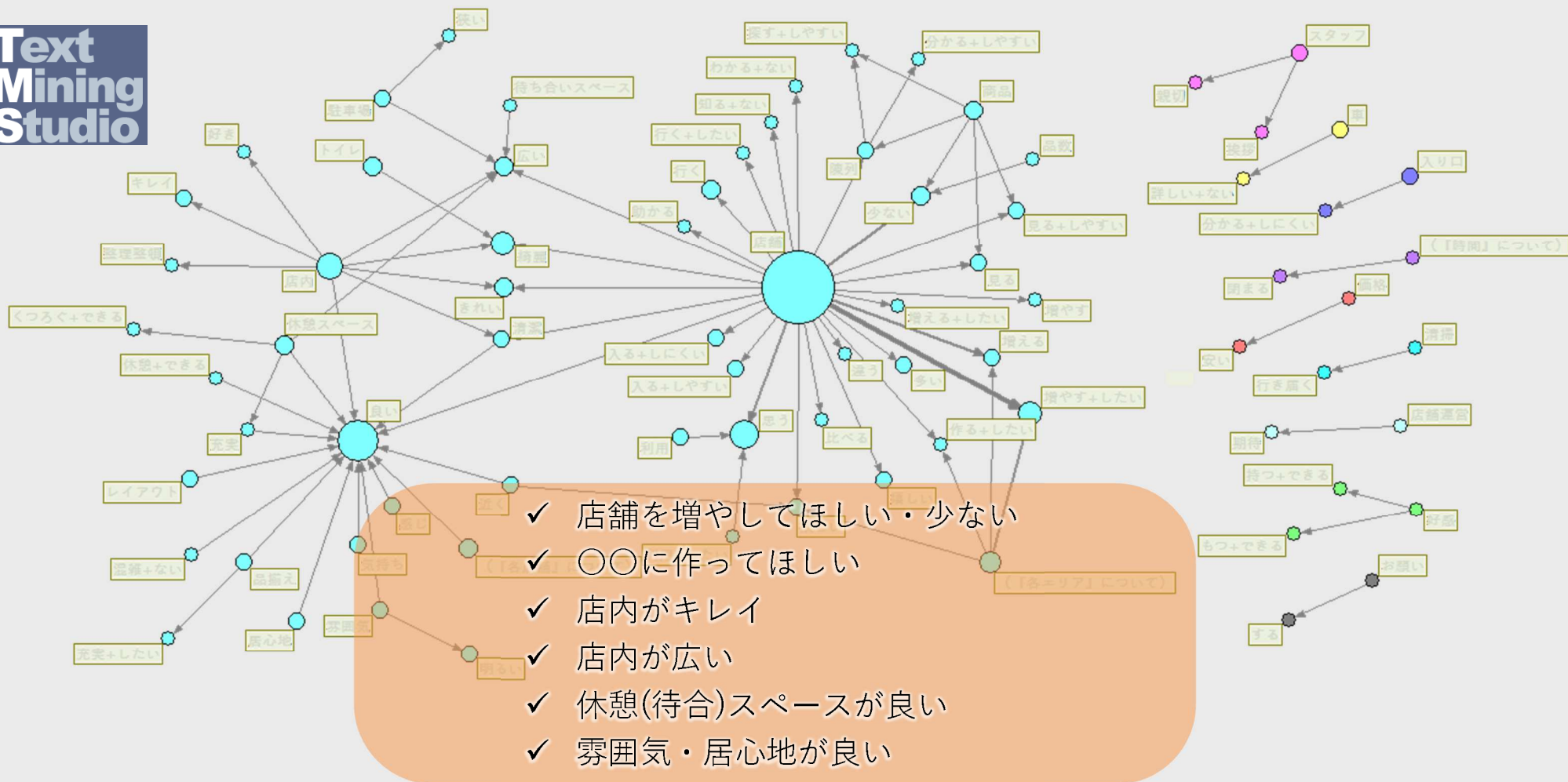
テキストマイニング手順

- [1] 分かち書き
- [2] 辞書の整理・作成
- [3] 集計・分析・可視化
- [4] 結果の読み込み



# 「店舗」「店内」について【係り受け】

Text Mining Studio



一歩踏み込んだ分析を行うため  
テーマを設定しますが **手間がかかります**



# Deep Learning

## 類義語辞書 作成編

**Deep  
Learner**

単語を数値（ベクトル）で表現させます

	態度表現付き置換語	A1	A2	A3	A4	A5	A6	A196	A197	A198	A199	A200
1	落とす	0.447	-0.152	-0.312	-0.495	0.359	0.127	0.180	0.153	-0.206	0.351	-0.189
2	その	0.186	-0.109	-0.164	0.041	0.128	0.192	0.212	-0.147	0.085	0.300	-0.206
3	水	0.398	-0.099	-0.317	-0.486	0.403	-0.006	0.305	-0.346	-0.144	0.490	-0.423
4	でも	0.357	-0.065	-0.109	-0.118	0.062	-0.119	0.073	-0.283	-0.106	0.236	-0.447
5	もっと	0.007	-0.042	-0.246	-0.763	0.646	0.277	0.250	-0.565	-0.844	0.375	-0.200
6	長い	0.262	-0.422	-0.176	-0.348	0.074	0.225	-0.157	0.015	0.223	0.341	-0.144
7	使う+したい	0.452	-0.019	-0.203	-0.342	0.395	-0.125	-0.195	0.142	0.010	0.296	-0.363
8	長持ち	-0.023	0.079	-0.067	-0.459	-0.022	0.005	0.260	0.151	-0.010	0.109	-0.273
9	携帯	0.672	-0.646	-0.392	-0.062	-0.020	0.681	-0.226	0.152	0.410	0.268	-0.215
10	欲しい	0.499	0.171	0.326	-0.960	0.011	-0.200	-0.079	-0.190	-0.285	0.971	-0.381
11	ラジオ	-0.011	-0.005	-0.073	-0.237	-0.147	0.159	-0.038	-0.147	-0.053	0.225	-0.160
12	聴く+できる	0.208	0.196	-0.357	-0.344	-0.205	-0.023	0.181	-0.014	-0.037	-0.120	-0.166
13	良い	0.606	0.632	-0.450	-0.363	0.216	-0.138	0.500	-0.054	0.122	0.653	0.523
14	電車	0.308	0.001	-0.064	-0.187	-0.069	0.471	0.012	-0.161	-0.075	-0.002	0.038
15	中	0.778	0.141	-0.189	-0.285	-0.050	0.611	-0.175	-0.449	-0.537	0.664	0.163
16	テレビ	0.191	0.267	-0.369	-0.554	0.071	0.299	0.036	-0.081	-0.148	0.480	-0.314
17	まず	0.196	0.050	-0.279	-0.233	-0.042	0.018	0.112	-0.198	-0.180	0.335	-0.330
18	わかる+ない	-0.060	0.025	-0.007	-0.178	-0.285	-0.104	0.136	-0.323	-0.257	-0.079	-0.122
19	最近	0.049	-0.310	0.011	-0.339	0.116	0.180	-0.099	-0.118	-0.046	0.393	-0.155
20	メール	0.153	0.387	-0.286	0.011	0.101	-0.248	0.375	-0.177	0.064	0.103	-0.432
21	読む	0.314	0.065	-0.450	-0.402	-0.094	0.190	0.265	-0.244	0.266	0.127	-0.594
22	文字	0.062	-0.060	-0.274	-0.451	0.247	0.043	0.059	0.061	0.072	-0.077	-0.550
23	大きい	0.196	0.155	-0.221	-0.238	0.220	0.191	0.329	0.260	0.229	-0.161	-0.431
24	表示	-0.006	0.077	-0.215	-0.196	0.068	-0.112	0.037	0.177	0.422	-0.329	-0.343
25	買い換える+したい	0.095	0.182	-0.206	-0.488	0.187	0.068	0.232	0.208	-0.071	0.190	-0.501
26	思う	0.186	0.143	-0.520	-0.913	-0.215	-0.148	-0.446	-0.375	-0.816	0.055	-0.630

## Deep Learner

The screenshot displays the 'Deep Learner Word Embedding' application interface. The main window is divided into several sections:

- モデル (Model):** Skip-Gram is selected.
- 学習パラメータ (Learning Parameters):** Includes options for approximation method (Negative Sampling), word count (5), Smoothing Exponent (0.75), embedding dimension (200), window size (5), mini-batch size (64), learning rate (0.01), and number of epochs (100).
- 環境設定 (Environment Settings):** Includes GPU usage and random seed options.
- 変数設定 (Variable Settings):** A table listing variables for training and validation.

列名	列属性	キー列	データ列
1	ファイルID	整数	
2	行ID		
3	文章ID		
4	単語ID		
5	見出し語		
6	原形		
7	置換語		
8	品詞		
9	品詞詳細		
10	係り先		
11	態度表現		
12	分かち書き		
13	態度表現作		
14	アンサンプル		

The 'Train & Validation' window shows a graph of Loss vs. Mini batches. The loss starts at approximately  $7.5 \times 10^1$  and decreases to about  $5.5 \times 10^1$  over 4,000 mini batches. The current epoch is 59. Below the graph are two circular progress indicators: 'train' (192 / 4582) and 'validation' (0 / 4582). A 'stop' button and a 'scores' button are also visible.

## 「商品」

	クエリ	類似語	類似度
1	商品	品数	0.849
2	商品	感じる	0.819
3	商品	品揃え	0.816
4	商品	比べる	0.806
5	商品	種類	0.805
6	商品	値段	0.802
7	商品	近く	0.800
8	商品	最近	0.800
9	商品		0.799
10	商品	店舗数	0.798
11	商品	もう少し	0.798
12	商品		0.796
13	商品		0.796
14	商品	多い	0.795
15	商品	他	0.795
16	商品	商品数	0.792
17	商品	他店	0.792
18	商品	店	0.790
19	商品	充実	0.786
20	商品	気がする	0.786

## 「店員」

	クエリ	類似語	類似度
1	店員	非常	0.970
2	店員	大変	0.970
3	店員	人	0.967
4	店員	買い物+できる	0.961
5	店員	すごい	0.961
6	店員	スタッフ	0.959
7	店員	接客態度	0.958
8	店員	声	0.958
9	店員	広々	0.956
10	店員	作業	0.955
11	店員	ゆったり	0.952
12	店員	気持ち	0.952
13	店員	入る	0.952
14	店員	休憩スペース	0.951
15	店員	元気	0.951
16	店員	スタッフさん	0.950
17	店員	従業員	0.950
18	店員	待合室	0.950
19	店員	キレイ	0.950
20	店員	店舗内	0.949

## 「綺麗」

	クエリ	類似語	類似度
1	綺麗	きれい	0.966
2	綺麗	清潔	0.951
3	綺麗	明るい	0.944
4	綺麗	広い	0.943
5	綺麗	入る+しやすい	0.941
6	綺麗	雰囲気	0.928
7	綺麗	キレイ	0.925
8	綺麗	見る+しやすい	0.923
9	綺麗	清潔感	0.916
10	綺麗	好感	0.914
11	綺麗	とても	0.911
12	綺麗	気持ち良い	0.909
13	綺麗	買い物+できる	0.909
14	綺麗	入る	0.907
15	綺麗	気持ち	0.907
16	綺麗	感じ	0.906
17	綺麗	持つ+できる	0.905
18	綺麗	居心地	0.903
19	綺麗	入る+しにくい	0.899
20	綺麗	広々	0.898

# Deep Learning

## テーマ設定編

## Deep Learner

## Seq2Seqモデルを使います

The screenshot displays the DeepLearner software interface. On the left, the 'モデルデザイン' (Model Design) section shows a vertical stack of layers: 出力層 (Output Layer), DecodeRNN層2, DecodeRNN層1, EncodeRNN層2, EncodeRNN層1, and 入力層 (Input Layer). In the center, the 'RNN層パラメーター一覧' (RNN Layer Parameter List) table provides details for each layer:

名前	出力次元数	タイプ
DecodeRNN層2	100	LSTM
DecodeRNN層1	100	LSTM
EncodeRNN層2	50	LSTM
EncodeRNN層1	100	LSTM

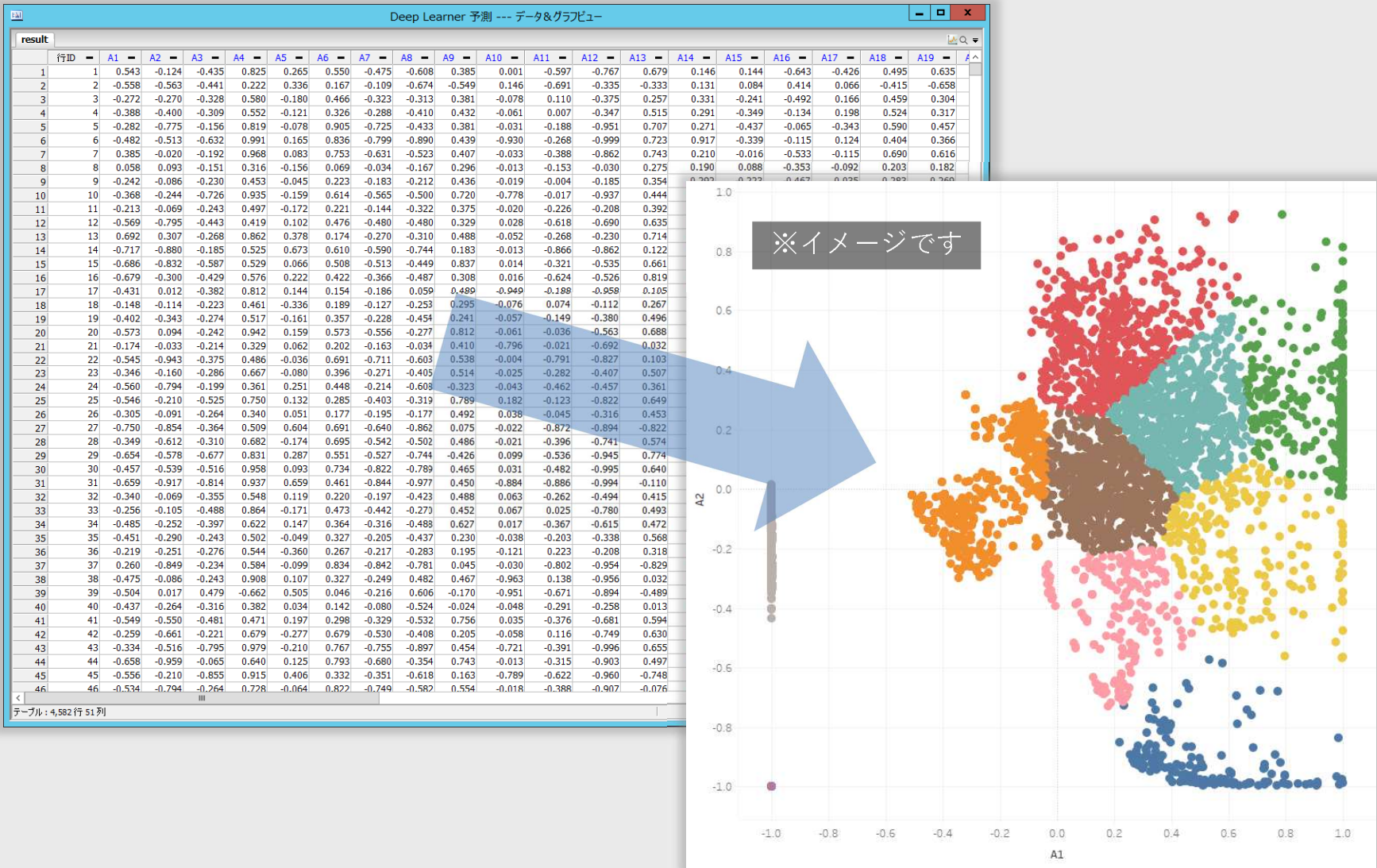
On the right, the 'Train & Validation' window shows a graph of Loss vs. Mini batches. The loss starts at approximately 5e+0 and decreases to about 1e+0 over 4,000 mini batches. Below the graph, the current epoch is 82. Two circular progress indicators show the number of training and validation samples: 1856 for training and 914 for validation.

# DeepLearnerでテーマ（特徴）を採る



Deep Learner

## 学習結果をクラスタリングします



# テキストマイニングスタジオの特徴表現 結果

Cluster-1	指標値
外装パーツ-取り寄せる	18.547
併設-運送	12.364

Cluster-2	指標値
スタッフ-良い	60.509
スタッフ-親切	51.036
対応-持つ+できる	36.927
良い-良い	32.49
スタッフ-対応	32.29
対応-対応	31.208
親切-親切	28.619
親切-親切	20.76
親切-親切	21.735
親切-親切	21.593
トイレ-綺麗	21.393
店内-良い	15.284
親切-良い	11.379
親切-良い	11.038
店内-広い	11.038
店内-広い	10.697
親切-知る+ない	10.697
元気-良い	9.332
お願い-申し上げる	8.79
挨拶-元気	8.79
対応-満足	8.79
店内-整理整頓	8.79
雰囲気-入る+しにくい	8.79

Cluster-3	指標値
雰囲気-良い	95.628
対応-良い	79.565
居心地-良い	63.152
サービス-良い	58.499
感じ-良い	53.591
休憩スペース-良い	42.436
店舗-思う	41.432
店舗-思う	39.919
店舗-思う	32.1
店舗-思う	22.819
店舗-思う	20.764
店舗-思う	19.266
充実-良い	18.915
店舗-良い	16.812
挨拶-良い	16.413
持ち合いスペース-広い	16.063
元気-良い	15.362
キッズスペース-良い	14.262
清潔感-良い	14.262

Cluster-4	指標値
価格-高い	87.699
価格-増やす+したい	42.846
価格-少ない	32.892
価格-少ない	32.833
価格-増やす+したい	30.886
価格-増やす+したい	29.902
価格-安い+したい	26.577
価格-安い+したい	21.905
価格-安い+したい	21.905
価格-安い+したい	14.617
価格-安い+したい	14.617
価格-安い+したい	14.597
価格-安い+したい	14.282
価格-増える	14.282
価格-少ない	13.948
イベント-安心+したい	11.961
イベント-安心+したい	11.961
イベント-増やす+したい	11.626
価格-安い	11.607

Cluster-5	指標値
【時間】(について)-営業+したい	44.572
【時間】-早い	35.729
【ラクター-商品-充実+したい	22.639
【カード-対応+したい	14.15
【時間】-早い+すぎる	13.443
現状-満足	32
現状-満足	32
現状-満足	6
現状-満足	10.966
現状-満足	66
現状-満足	13
現状-満足	3
現状-満足	8.49
現状-満足	8.49
現状-満足	8.49
現状-満足	8.136
現状-満足	8.136
現状-満足	8.136
現状-満足	8.136
現状-満足	8.136



スタッフ

店内

良い

価格

営業時間

広告・イベント

Cluster-6	指標値
店舗-増やす+したい	132.897
店舗-少ない	85.034
店舗-きれい	58.424
店舗-綺麗	54.319
店舗-増える	32.839
店舗-広い	22.725
店舗-入る+しやすい	17.74
店舗-陳列	13.635
店舗-陳列	13.484
店舗-陳列	13.415
店舗-陳列	13.415
店舗-陳列	10.775
店舗-良い	9.234
店舗-充実	9.09
店舗-気に入る	9.09
店舗-よる	9.09
店舗-速い	9.09
店舗-少ない-良い	8.87
店舗-入る+しにくい	8.65
店舗-明るい	8.65

Cluster-7	指標値
店舗-増やす+したい	197.223
【具体的な地域】-増やす+したい	149.124
【具体的な地域】-出店+したい	77.37
店舗-増える	40.811
【具体的な地域】-作る+したい	39.594
店舗-作る+したい	30.796
【具体的な地域】-欲しい	26.396
店舗-欲しい	15.942
店舗-欲しい	15.987
店舗-欲しい	15.997
店舗-欲しい	21.77
店舗-助かる	17.905
【具体的な地域】-つくる+したい	13.198
【具体的な地域】-良い	13.198
店舗-良い	10.243
【具体的な地域】-ネットチェック導入	8.799
もの-欲しい	8.799
ポイント-欲しい	8.799
【具体的な地域】-行く+したい	8.799

Cluster-8	指標値
店舗-増やす+したい	146.663
店舗-良い	99.132
店舗-少ない	82.851
店舗-少ない	76.811
店舗-増やす+したい	36.681
店舗-増やす+したい	36.097
店舗-増やす+したい	29.411
商品-少ない	21.008
商品-少ない	14.483
商品-少ない	10.22
商品-少ない	10.382
商品-少ない	17.726
商品-安い+したい	14.968
商品-増やす	14.968
種類-増やす+したい	14.706
種類-増やす+したい	11.817
商品-見る+しやすい	11.554
商品-中途半端	11.423

Cluster-9	指標値
ポイント-増える	7.607
利用+したい-思う	7.395
ウォーターサーバー-置く	6.94
車-持つ	6.94
スペース-置く	5.552
場所-嬉しい	5.552
機会-増える	4.778
車-嬉しい+ない	4.778
【時間】(について)-営業	4.164
【時間】(について)-置く	4.164
【時間】(について)-閉まる	4.164
ポイント-貯まる	4.164
ポイント-貯める	4.164
購入+したい-思う	4.164
時間-かかる	4.164
入-分かる+しやすい	4.164
店-運ぶ	4.164
女性-入る+しやすい	3.443
店舗-見る	3.443
買い物-行く	3.443

Cluster-10	指標値
ポイント-する	175.182
ポイント-使う+までほしい	76.064
ポイント-続ける+したい	66.891
00ポイント-プレゼント	25.952
ポイント-貯まる	25.281
ポイント-増大+したい	22.372
ポイント-チェック-お願い	21.478
ポイント-お願い	19.97
ポイント-お願い	17.45
ポイント-嬉しい	17.45
ポイント-サービス-続ける+したい	13.423
ポイント-チェック-素晴らしい	13.423
ポイント-貯める	13.423
00ポイント-続ける+したい	13.199
ポイント-行く	13.199

品揃え

商品

店舗 (感想)

店舗 (要望)

ポイント