



「ID-POS分析はAIで進化する」
最新事例と実践活用の課題

Visual Mining Studio(PLSA)、
BayoLinkS、DeepLearnerの実践活用方法



株式会社 IDプラスアイ

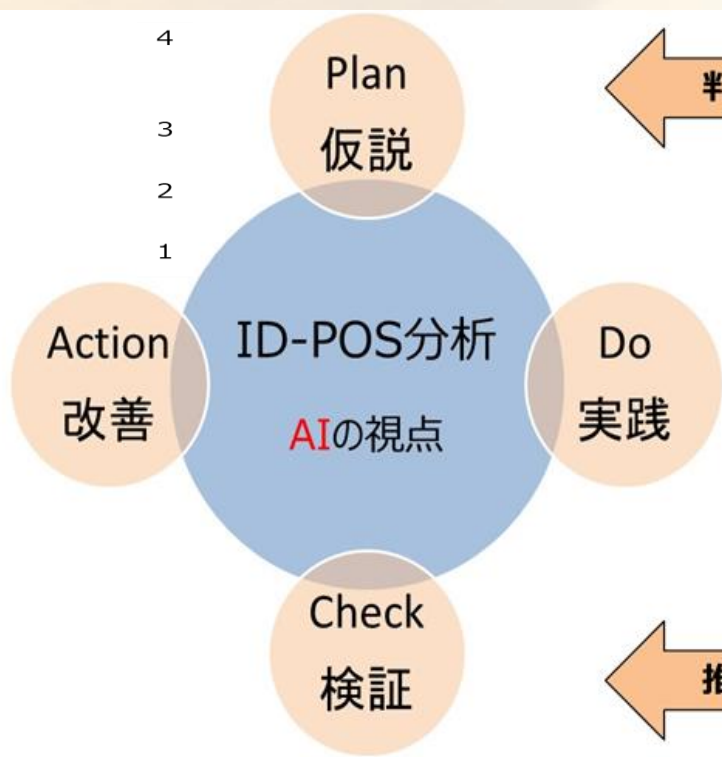
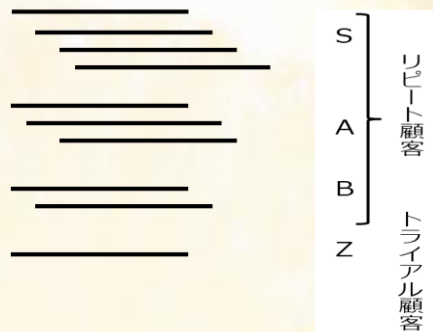


最新事例公開

AIの目的は、分析ではなく、「判別、予測、推論」にあり！

POS分析
 延べ客数 (客数) = **ID-POS分析**
 ユニーク客数 (ID客数) × 頻度 (客数/ID客数)

10 = 4 × 2.5



判別のAI : **PLSA**

- PLSA (Probabilistic Latent Semantic Analysis : 確率的潜在意味解析)
- 商品と顧客を同時に分析し、確率的(ソフト) にクラスターを作り上げるAIのアルゴリズム

予測のAI : **DL**

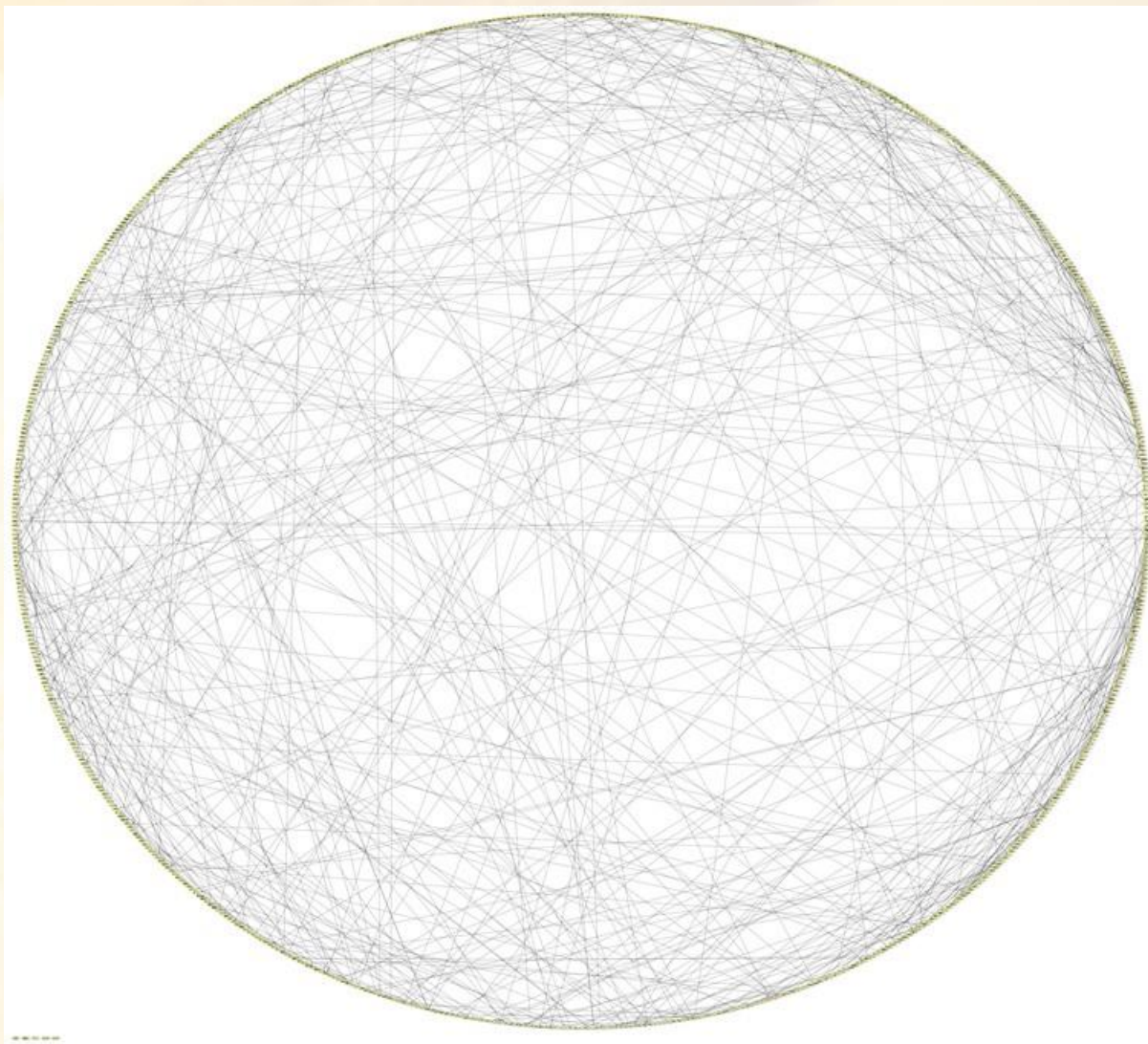
- DL (Deep Learning : 深層学習)
- 実践に向けて、需要予測をするためのAIのアルゴリズム
- 数値予測や顧客、商品スコアの推論に適している。

推論のAI : **BN**

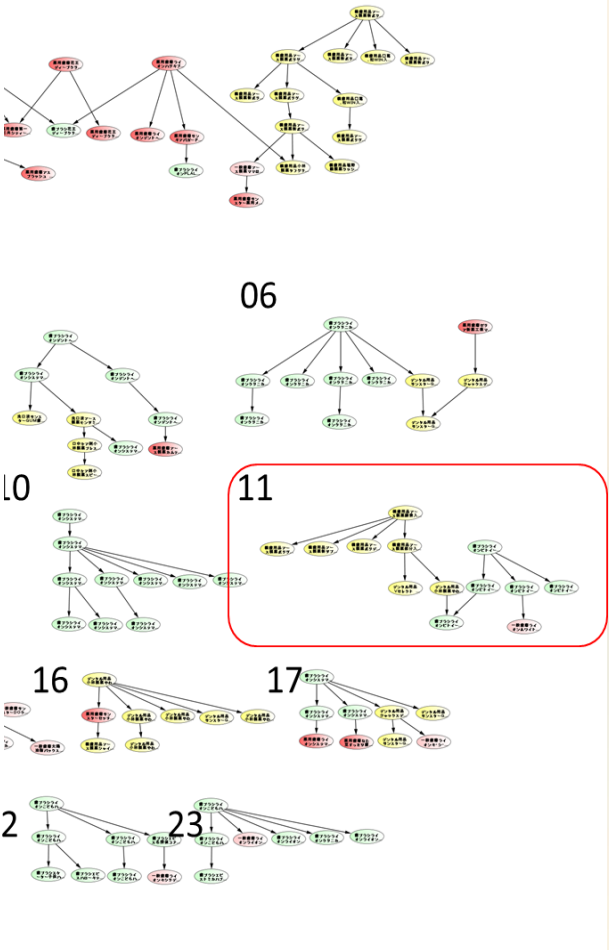
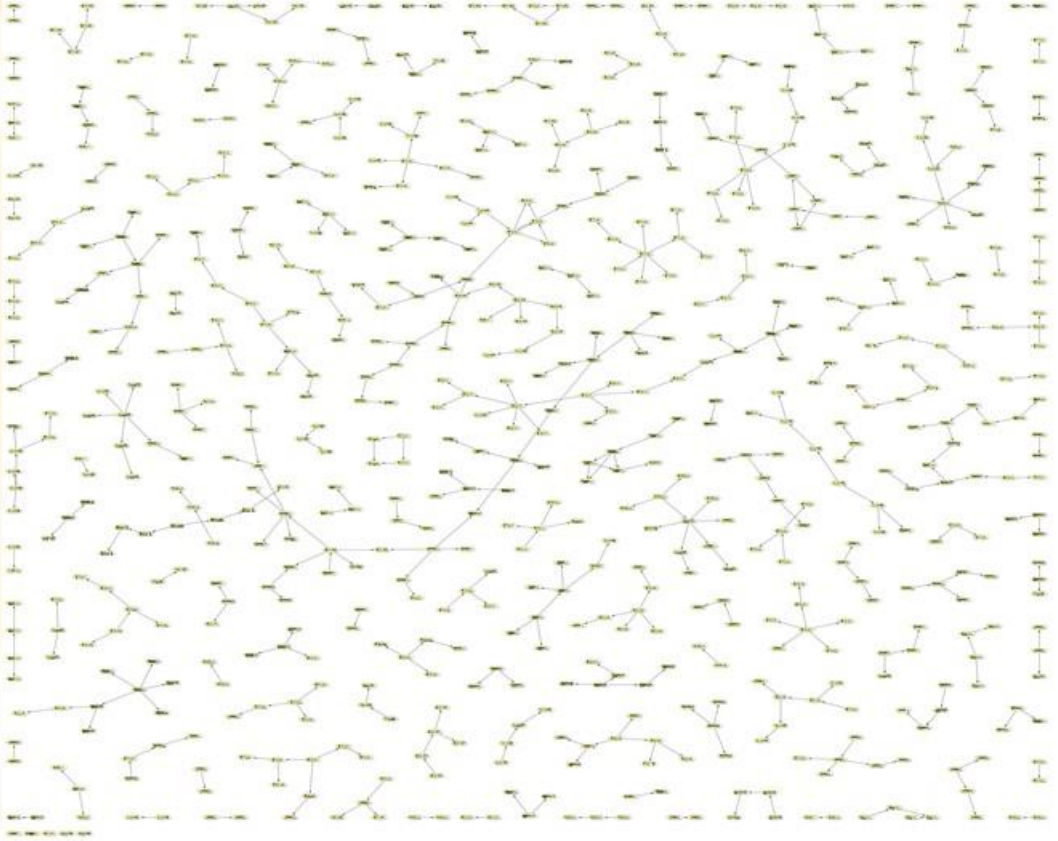
- BN (Bayesian Network : ベイジアンネットワーク)
- 実践した結果得られた学習データをもとに、様々な推論を行い、仮説の改善につなげるためのAIのアルゴリズム
- PLSAと相性がよく、組み合わせて活用すると、より精度の高い推論が可能

AIの目的は、人間の知能、判別力、予測能力、推論力を支援し、高める！

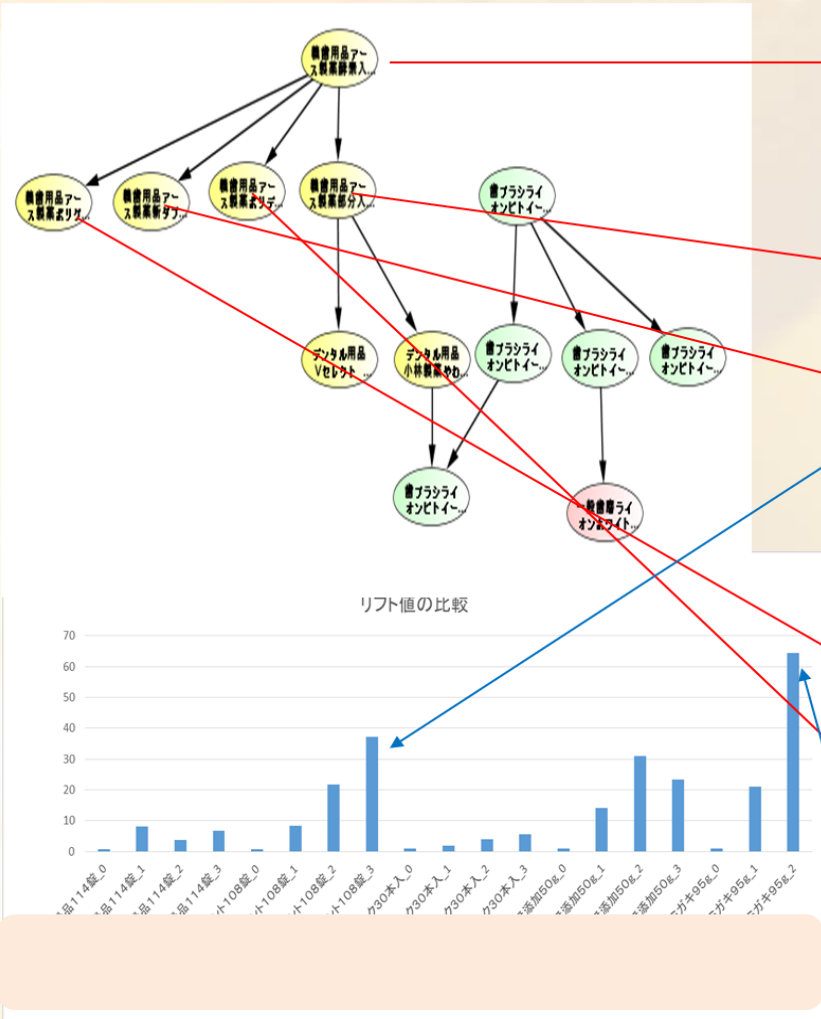
1.事例_01 (推論) : Bayesian Network と ID-POS分析



AI棚割り分析事例：親子関係の推論



AI棚割り分析事例：親子関係の推論



	state	value	value	evidence?
5	品114錠	0	96%	0% FALSE
6	品114錠	1	3%	0% FALSE
7	品114錠	2	1%	0% FALSE
8	品114錠	3	0%	100% TRUE
9	品114錠_0	0	95%	68% 1
10	品114錠_1	1	3%	25% 8
11	品114錠_2	2	1%	4% 4
12	品114錠_3	3	1%	4% 7
113	錠_0	0	98%	75% 1
114	錠_1	1	1%	11% 8
115	錠_2	2	0%	7% 22
116	錠_3	3	0%	7% 37
472	本入_0	0	99%	98% 1
473	本入_1	1	0%	1% 2
474	本入_2	2	0%	0% 4
475	本入_3	3	0%	1% 5
500	ウダー無添加50g_0	0	99%	89% 1
501	ウダー無添加50g_1	1	0%	4% 14
502	ウダー無添加50g_2	2	0%	4% 31
503	ウダー無添加50g_3	3	0%	4% 23
674	歯のハミガキ95g_0	0	100%	89% 1
675	歯のハミガキ95g_1	1	0%	4% 21
676	歯のハミガキ95g_2	2	0%	7% 64

AI棚割り分析事例：学習データ

商品：約500

顧客
5,000人

会員番号	10月	大然水南	ケース	おいしい	ケース	大然	おいしい	水	ナノ	素水	ヨー	クリー	もも	いろ	ワイ	ッテル	いろ	は	コ	ント	レック	朝	換	み	オレ	伊	藤
2.82E+12	3	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.87E+12	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2.82E+12	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

AI棚割り分析の手順：

1. 分析対象を決める
 - 棚の分析が基本だが、プラスアルファを検討
 - 周辺の商品を加えるか、クロスMDの商品を加えるか？
2. 分析期間を決める
 - 原則6ヶ月以上は欲しい。
 - 頻度分析が基本なので、レポートが十分に発生する期間が必須
 - 頻度を0+SA(ロイヤル) +B+Z(トライアル) の4つを基本とする。
3. BNで因果推論を実施し、構成クラスターを確認
 - 主要クラスターをもとに棚割りの骨格を再検討
 - 主要クラスターの品揃えを検討
 - 主要クラスターの販促を検討

親子関係のAIでの評価ポイント：

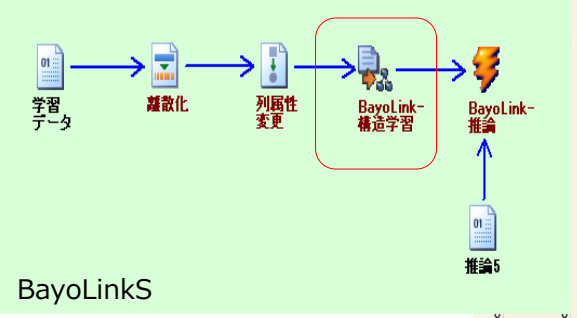
- (1) ある商品について、残りの商品から適当なものを選択しそれを親とする。
- (2) 親とのリンク(ツリー)を作り、評価値を計算する。
- (3) 他の組合のツリーの評価値も計算する。その中から最も評価値のよい親を採用する。
- (4) **全ての商品**に対して (1) ~ (3) を行う。全商品のツリーを併せて**1つのモデル**とする。
- (5) ネットワークに双方向リンクや**循環があれば、これら回避するために一部のリンクを削除する。**

AIC - 赤池の情報量

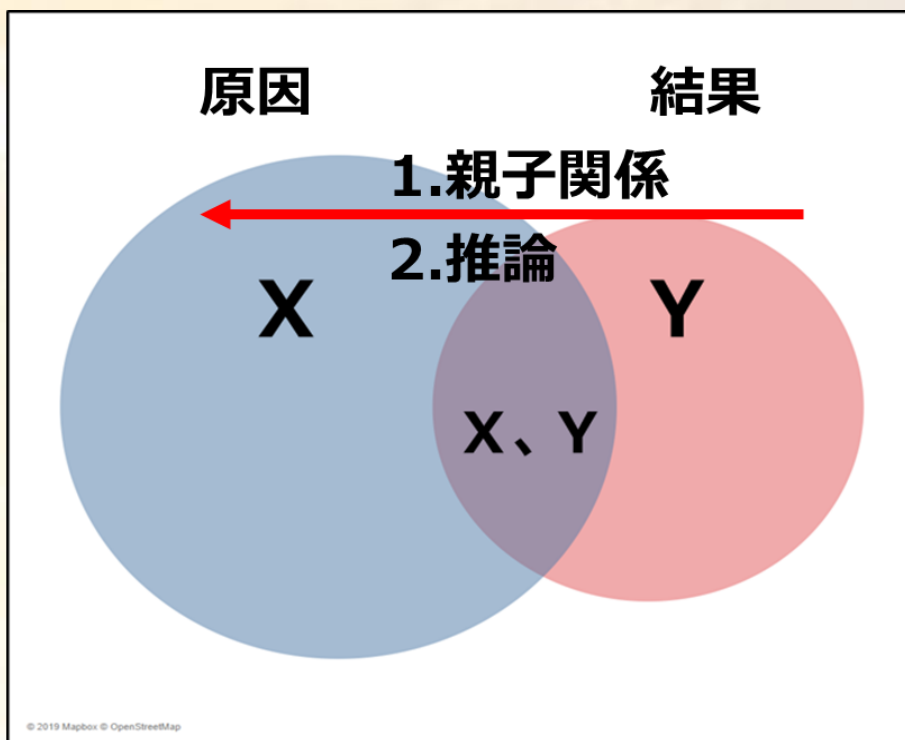
$$AIC = -2 \left(ML - \sum_{i=1}^n (r_i - 1) s_i \right)$$

ID-POS分析のF(頻度) の研究成果(離散化) を入れて学習データを作成

- 1回、2回、3回以上 (SABZの視点) を導入



参考：超併売（ID-POS分析のAIでの進化）



ベイズの定理

$$P(X|Y) = P(Y|X) \frac{P(X)}{P(Y)}$$

逆方向 順方向

X = 原因、Y = 結果

* 結果を原因から推論できる。

↓

時間とすると、因果逆転の計算が可能

↓

医療診断、故障、仮説検証（PDCA）

ID-POSはAIで進化する

- ユニーク客数、頻度の視点

(ID-POS分析特有の視点)

- ユニーク客数ごとの購入商品
- 頻度での関係性分析
 - SABZ（頻度区分）
 - リピート、トライアル、..

- 超併売：AI（ベイジアンネットワーク）

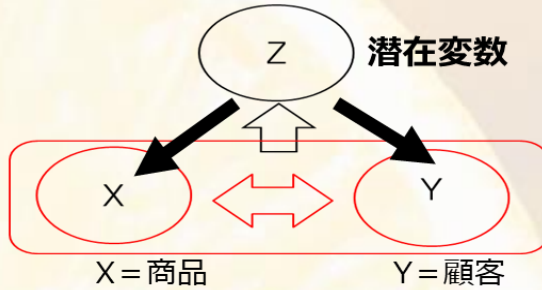
- 因果関係図（親子）を構築
- 因果推論を実施

2.事例_02 (判別) : PLSAと ID-POS分析

PLSA : ソフトクラスタリング
 Probabilistic Latent Semantic Analysis

確率的 潜在 意味 解析

$$\hat{P}(X_n, Y_m) = \sum_{k=1}^L P(Z_k)P(X_n | Z_k)P(Y_m | Z_k)$$



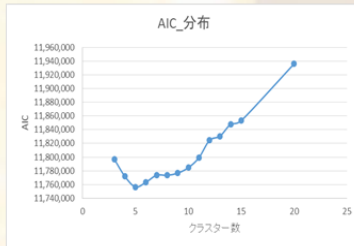
AIC - 赤池の情報量

$$AIC = -2 \left(ML - \sum_{i=1}^p (r_i - 1) s_i \right)$$

赤池情報量規準:

ウィキペディア (Wikipedia)

- 元々は An Information Criterion, のちに Akaike's Information Criterion と呼ばれるようになる)は、統計モデルの良さを評価するための指標である。単にAICとも呼ばれ、この呼び方のほうが一般的である。統計学の世界では非常に有名な指標であり、多くの統計ソフトに備わっている。元統計数理研究所所長の赤池弘次が1971年に考案し1973年に発表した。



AI仮想クラスターの参考事例(青果) :

Z1クラスター

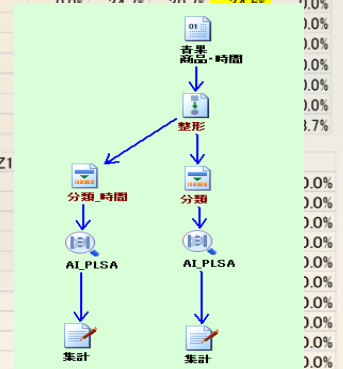
商品名	Z1	Z2	Z3	Z4	Z5	Z1	Z2	Z3	Z4	Z5
1 シヤインマスカット	10.3%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
2 もやし500	10.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
3 グレープフルーツR	8.3%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
4 オニオンサラダA	6.1%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
5 キャベツミックス	6.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
6 たまねぎA	5.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
7 ブロッコリーS	3.6%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
8 キウイフルーツPP	3.5%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
9 たねなしぶどうA	3.3%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
10 たまねぎサラダ	2.9%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%

会員番号	Z1	Z2	Z3	Z4	Z5	Z1	Z2	Z3	Z4	Z5
1 2.8159E+12	0.3%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
2 2.8159E+12	0.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
3 2.8159E+12	0.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
4 2.833E+12	0.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
5 2.8159E+12	0.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
6 2.8159E+12	0.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
7 2.8159E+12	0.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
8 2.8159E+12	0.2%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
9 2.8159E+12	0.2%	0.0%	0.0%	0.0%	0.0%	99.1%	0.0%	0.9%	0.0%	0.0%
10 2.8159E+12	0.2%	0.0%	0.0%	0.0%	0.0%	88.9%	11.1%	0.0%	0.0%	0.0%

Z4クラスター

商品名	Z1	Z2	Z3	Z4	Z5	Z1	Z2	Z3	Z4	Z5
1 もやし200	0.0%	0.0%	0.0%	32.8%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%
2 ぶなしめじ	0.0%	4.8%	0.0%	19.7%	0.0%	0.0%	18.7%	0.0%	81.3%	0.0%
3 トマトF	0.0%	0.0%	0.0%	15.4%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%
4 えのき茸A	0.0%	7.5%	12.7%	7.1%	0.0%	0.0%	24.7%	20.7%	24.6%	0.0%
5 エリンギ	0.0%	5.4%	0.0%	5.2%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%
6 にんじん	0.0%	5.6%	3.4%	4.9%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%
7 ミックスサラダ	0.0%	0.0%	0.0%	4.5%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%
8 細ごぼう	0.0%	0.0%	0.0%	2.6%	0.0%	0.0%	0.0%	0.0%	0.0%	1.0%
9 いら	0.0%	2.5%	0.0%	1.9%	0.0%	0.0%	0.0%	0.0%	0.0%	1.7%
10 しょうがK	0.0%	1.2%	0.0%	1.6%	1.6%	0.0%	0.0%	0.0%	0.0%	1.0%

会員番号	Z1	Z2	Z3	Z4	Z5	Z1
1 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
2 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
3 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
4 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
5 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
6 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
7 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
8 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
9 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%
10 2.8159E+12	0.0%	0.0%	0.0%	0.1%	0.0%	0.0%



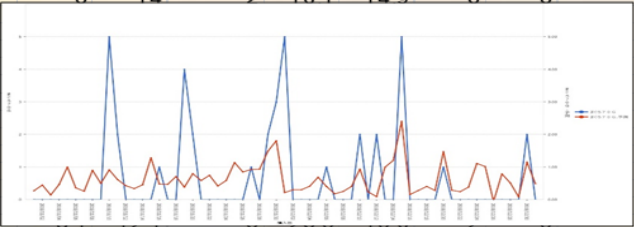
Visual Mining Studioの機能 "二項ソフトクラスタリング" を利用

様々な要素による組み合わせでの学習データ

1	2	3	4	5	6	7	8		
顧客ID	性別	年代	郵便番号	商品名	小分類名	購入日	購入時間	数量	
280846	男性	70	1820000	マジックリン500ML	換気扇レンジ	2018年5月31日	90800	1	
280846	男性	70	1820000	マジックリン500ML	換気扇レンジ	2018年7月26日	100100	2	
280846	男性	70	1820000	マジックリン500ML	換気扇レンジ	2018年10月9日	92600	2	
280846	男性	70	1820000	水切りストッキングA30枚	水切袋	2018年4月10日	103200	1	
280846	男性	70	1820000	水切りストッキングA30枚	水切袋	2018年5月20日	94100	1	
280846	男性	70	1820000	水切りストッキングB30枚	水切袋	2018年2月6日	90700	1	
2		1	2	3	4	5	6	7	8
2	X	顧客ID	性別	年代	郵便番号	商品名	小分類名	購入日	購入時間
2	28	7	6	5	4	3	2	1	0
2	Y	性別							
2		年代	年代						
2		郵便番号	郵便番号	郵便番号					
2		商品名	商品名	商品名	商品名				
2		小分類名	小分類名	小分類名	小分類名	小分類名			
2		購入日	購入日	購入日	購入日	購入日	購入日		
2		購入時間	購入時間	購入時間	購入時間	購入時間	購入時間	購入時間	
200640	男性	70	1820000	ポリデント72錠	義歯用剤	2018年4月12日	92700	1	
280846	男性	70	1820000	ポリデント72錠	義歯用剤	2018年4月22日	100000	1	
280846	男性	70	1820000	ポリデント72錠	義歯用剤	2018年7月5日	92900	1	
280846	男性	70	1820000	ポリデント72錠	義歯用剤	2018年9月1日	105600	1	
280846	男性	70	1820000	ポリデント72錠	義歯用剤	2018年11月15日	100500	1	
280846	男性	70	1820000	新ポリグリッパ40	義歯用剤	2018年12月25日	91500	1	
280846	男性	70	1820000	新ポリグリッパ75	義歯用剤	2018年3月12日	163400	1	
280846	男性	70	1820000	新ポリグリッパTP75	義歯用剤	2018年10月9日	92600	1	
280846	男性	70	1820000	ナイロンスポンジ5P	たわし	2018年7月26日	100100	1	
280846	男性	70	1820000	キッチンハイター600	キッチン漂白剤	2018年6月1日	90500	1	
280846	男性	70	1820000	キッチンハイター600	キッチン漂白剤	2018年8月25日	91700	1	
280846	男性	70	1820000	マジックリンスプレ400	換気扇レンジ	2018年1月9日	161500	1	
280846	男性	70	1820000	マジックリンスプレ400	換気扇レンジ	2018年1月12日	100100	1	
280846	男性	70	1820000	マジックリンスプレ400	換気扇レンジ	2018年3月7日	90900	2	
280846	男性	70	1820000	マジックリンスプレ400	換気扇レンジ	2018年5月2日	91300	1	
280846	男性	70	1820000	マジックリンスプレ400	換気扇レンジ	2018年11月25日	94100	1	

3.事例_03 (予測) : Deep Learning と ID-POS分析

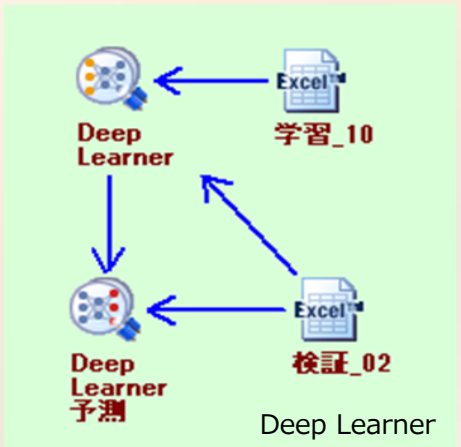
購入日	曜日	日照	最低	降水量	最高	平均	昨年	先週	ポイント	まぐろ70G	38	38	100%
2018/11/1	木	5.5	11.7	0	19.9	14.8	0	0	0	1	0	0	0%
2018/11/2	金	10.3	11	0	19.1	14.9	0	0	0	1	0	0	0%
2018/11/3	土	4.6	11.4	0	19.9	15.6	0	0	0	1	0	0	0%
2018/11/4	日	0	14	2	16.1	14.9	0	0	0	1	0	0	0%
2018/11/5	月									5	0	1	0%
2018/11/6	火									1	0	0	0%
2018/11/7	水									1	0	0	0%
2018/11/8	木									1	0	1	0%
2018/11/9	金									1	0	0	0%
2018/11/10	土									5	5	1	549%
2018/11/11	日									1	2	1	326%
2018/11/12	月									1	0	0	0%
2018/11/13	火									1	0	0	0%
2018/11/14	水									1	0	0	0%
2018/11/15	木									5	0	1	0%
2018/11/16	金									1	1	0	212%
2018/11/17	土									1	0	0	0%
2018/11/18	日									1	0	1	0%
2018/11/19	月									1	4	0	0%
2018/11/20	火									5	2	1	251%
2018/11/21	水									1	0	1	0%
2018/11/22	木									1	0	1	0%
2018/11/23	金									1	0	0	0%
2018/11/24	土	3.3	8.5	0	13.7	10.6	2	0	0	1	0	1	0%
2018/11/25	日	8.7	7	0	15.3	11	3	0	0	5	0	1	0%
2018/11/26	月	6	8.7	0	16.8	13.1	5	4	1	1	0	1	0%
2018/11/27	火	5.4	10.6	2.5	17.9	14.1	5	2	1	1	1	1	108%
2018/11/28	水	7.2	10.9	0	17.6	14.7	5	0	1	1	0	1	0%
2018/11/29	木	5.2	11.9	0.5	15.9	13.4	9	0	1	2	1	1	135%
2018/11/30	金	8.6	9.6	0	16	13	10	0	5	3	2	1	166%
2018/12/1	土	8.4	8.5	0	18.3	12.6	0	0	1	5	0	0	0%
2018/12/2	日	0.3	8.8	0	12	10.6	0	0	1	1	0	0	0%
2018/12/3	月	1.4	8.9	0	15.7	13.1	1	0	1	0	0	0	0%
2018/12/4	火	7.8	12	15.5	23.3	17.9	1	1	1	1	0	0	0%
2018/12/5	水	4.2	14.6	0	20.5	17.6	0	0	5	1	0	1	0%
2018/12/6	木	0	8.8	18.5	14.7	10.7	0	2	1	1	0	0	251%
2018/12/7	金	4.2	11	0	18.2	13.6	0	3	1	1	0	0	0%
2018/12/8	土	1.5	9.3	0	13	10.4	0	5	1	1	0	0	0%
2018/12/9	日	0.8	6	0	10.3	8.2	1	0	1	1	0	0	0%
2018/12/10	月	0.8	5.6	0	9.6	7	2	0	5	2	1	1	214%
2018/12/11	火	0.6	3.1	6.5	7.3	5.1	0	0	1	1	0	0	0%
2018/12/12	水	0	3.6	19.5	10.9	7.4	0	0	1	2	0	0	0%
2018/12/13	木	1	5.9	0	10.1	7.9	4	1	1	0	1	0	0%
2018/12/14	金	6.6	4	0	10.5	6.9	5	0	1	1	0	1	0%
2018/12/15	土	4.7	4.7	0	10.1	6.6	11	0	5	5	2	1	208%
2018/12/16	日	1.5	2	1	6.9	4.5	0	0	1	1	0	0	0%
2018/12/17	月	4.1	4.3	5	12	7.7	0	2	1	1	0	0	0%
2018/12/18	火	8	6.3	0	13.4	9.4	0	0	1	1	0	0	0%
2018/12/19	水	9.2	4.7	0	13.8	9.5	0	2	1	1	0	0	0%
2018/12/20	木	6.7	6.4	0	14	9.9	6	0	5	1	1	1	68%
2018/12/21	金	9.1	5.4	0	14	9.9	0	0	1	1	0	0	0%
2018/12/22	土	0	8.9	1.5	12	10.2	0	5	1	1	0	0	0%
2018/12/23	日	0.2	9.4	2.5	12.7	10.7	0	0	1	1	0	0	0%
2018/12/24	月	2.8	5.4	0	11.5	8.5	5	0	1	1	0	1	0%
2018/12/25	火	8.9	4.4	0	10	6.9	1	0	5	1	0	1	0%
2018/12/26	水	3.7	3.2	0	12	8.1	0	0	1	1	0	(O)	0%
2018/12/27	木	7.7	6	0	13.9	9.5	3	1	1	1	0	1	0%
2018/12/28	金	7.2	2.9	0	8.8	5.4	1	0	1	1	0	1	0%
2018/12/29	土	7.5	1.1	0	10.7	5.2	0	0	1	1	0	0	0%
2018/12/30	日	6.8	2.4	0	10.4	5.5	3	0	5	2	1	1	174%
2018/12/31	月	5.6	2.9	0	8.9	6	1	0	1	1	0	0	0%



2ヶ月:61日分をDLで予測(10ヶ月の学習データ)
 * 週ごとの跳幅の予測への対応が十分でない。
 * 全体としては38/38で100%の予測
 * 1個以上誤差日が5日/61日=約8%
 * 92%が1個以下の誤差、0予測も多数。

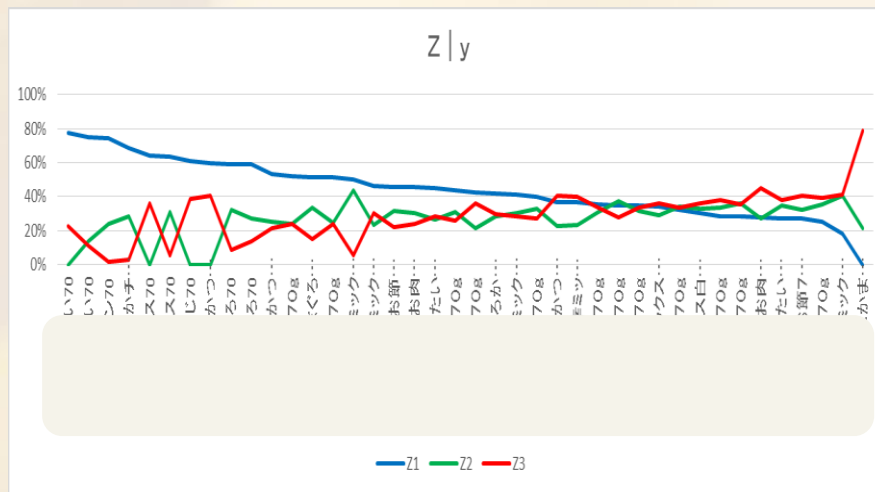
サンプル分析：まぐろ70G

- 気象データと実績値+昨年データ、先週データ+ポイントを説明変数とした。
- 1個/日の商品の需要予測であり、sell 1 buy 1でも良いが、..
- 週1回、はねる時の在庫の確保が課題

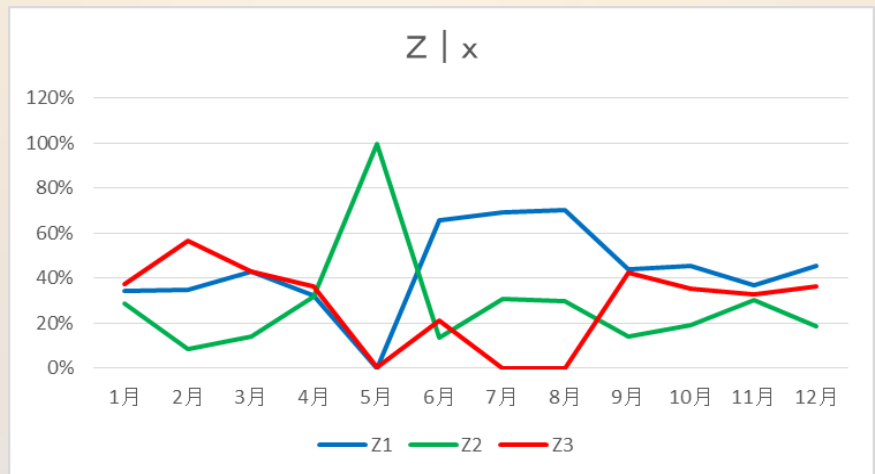


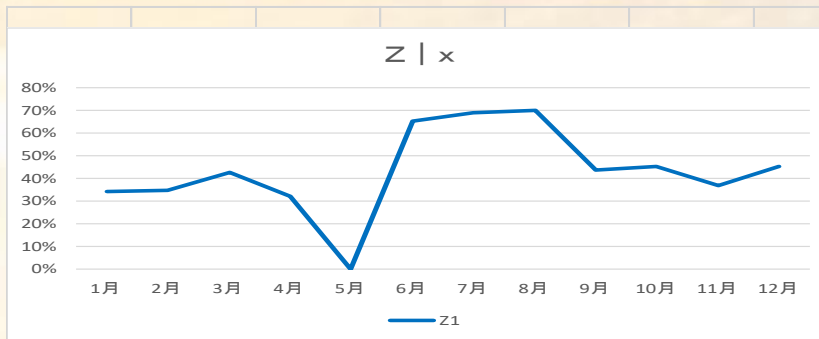
参考_01 : PLSAでクラスタリング

商品名	Z1	Z2	Z3	Z1	Z2	Z3
1 子猫まぐろたい70	77%	0%	23%	3%	0%	1%
2 老猫まぐろたい70	75%	14%	11%	6%	2%	1%
3 チキン70	75%	24%	2%	5%	3%	0%
4 子猫やわらかチキン70	69%	28%	3%	4%	2%	0%
5 魚肉ミックス70	64%	0%	36%	4%	0%	3%
6 子猫魚ミックス70	63%	31%	6%	3%	3%	0%
7 まぐろあじ70	61%	0%	39%	3%	0%	3%
8 老猫しらすかつお70	60%	0%	40%	3%	0%	3%
9 老猫まぐろ70	59%	32%	9%	5%	4%	1%
10 まぐろ70	59%	27%	14%	3%	2%	1%

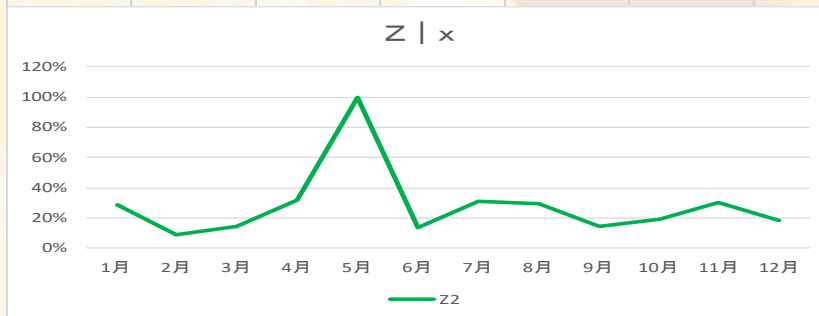


購入月	Z1	Z2	Z3	Z1	Z2	Z3
1月	34%	29%	37%	7%	10%	12%
2月	35%	9%	57%	7%	3%	17%
3月	43%	14%	43%	7%	4%	12%
4月	32%	32%	36%	5%	9%	9%
5月	0%	100%	0%	0%	25%	0%
6月	66%	13%	21%	12%	4%	6%
7月	69%	31%	0%	14%	10%	0%
8月	70%	30%	0%	13%	9%	0%
9月	44%	14%	42%	9%	4%	13%
10月	46%	19%	35%	8%	6%	10%
11月	37%	30%	33%	7%	10%	10%
12月	45%	18%	36%	9%	6%	11%

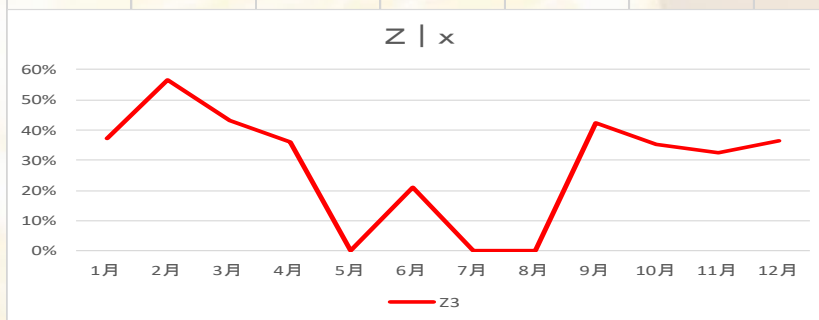




商品名	Z1	Z2	Z3	Z1	Z2	Z3
1 子猫まぐろたい70	77%	0%	23%	3%	0%	1%
2 老猫まぐろたい70	75%	14%	11%	6%	2%	1%
3 チキン70	75%	24%	2%	5%	3%	0%
4 子猫やわらかチキン70	69%	28%	3%	4%	2%	0%
5 魚肉ミックス70	64%	0%	36%	4%	0%	3%



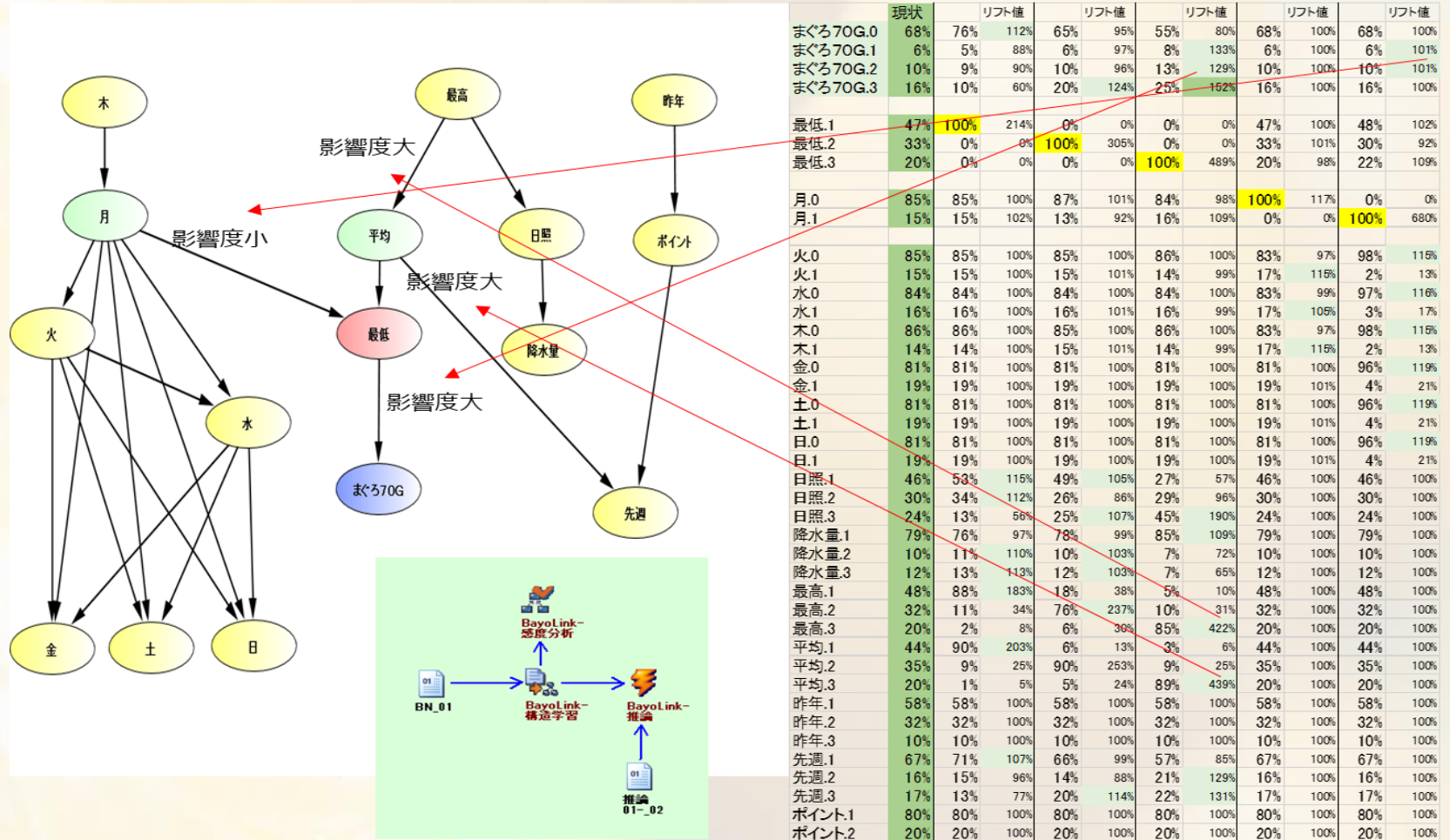
商品名	Z1	Z2	Z3	Z1	Z2	Z3
1 子猫魚ミックス70	50%	44%	6%	3%	4%	1%
2 老猫魚ミックス70	18%	40%	41%	1%	5%	4%
3 まぐろ白身魚70	35%	37%	28%	2%	3%	2%
4 老猫まぐろ70	29%	36%	35%	2%	4%	4%
5 子猫まぐろ70	25%	36%	39%	2%	4%	4%



商品名	Z1	Z2	Z3	Z1	Z2	Z3
1 子猫かにかま70	0%	21%	79%	0%	2%	6%
2 老猫魚肉ミックス70	28%	27%	45%	2%	3%	5%
3 老猫魚ミックス70	18%	40%	41%	1%	5%	4%
4 老猫かつお節70	27%	32%	41%	2%	3%	4%
5 老猫しらすかつお70	60%	0%	40%	3%	0%	3%

参考_02 : BNで説明変数の推論

気象データの推論





実戦活用の課題

2. AIからのアウトプットをID-POS分析に組み込む際の課題



商品DNA、顧客DNAのポイント：

- 商品DNAの「顧客」の場合
 - 顧客属性（性別、年代、職業、年収、…）
 - 購入頻度、購入金額、購入点数、…
 - 顧客DNAを援用
- 顧客DNAの「商品」の場合
 - 商品属性（容量、価格、メーカー、…）
 - 購入頻度、購入金額、購入点数
 - 商品DNAを援用

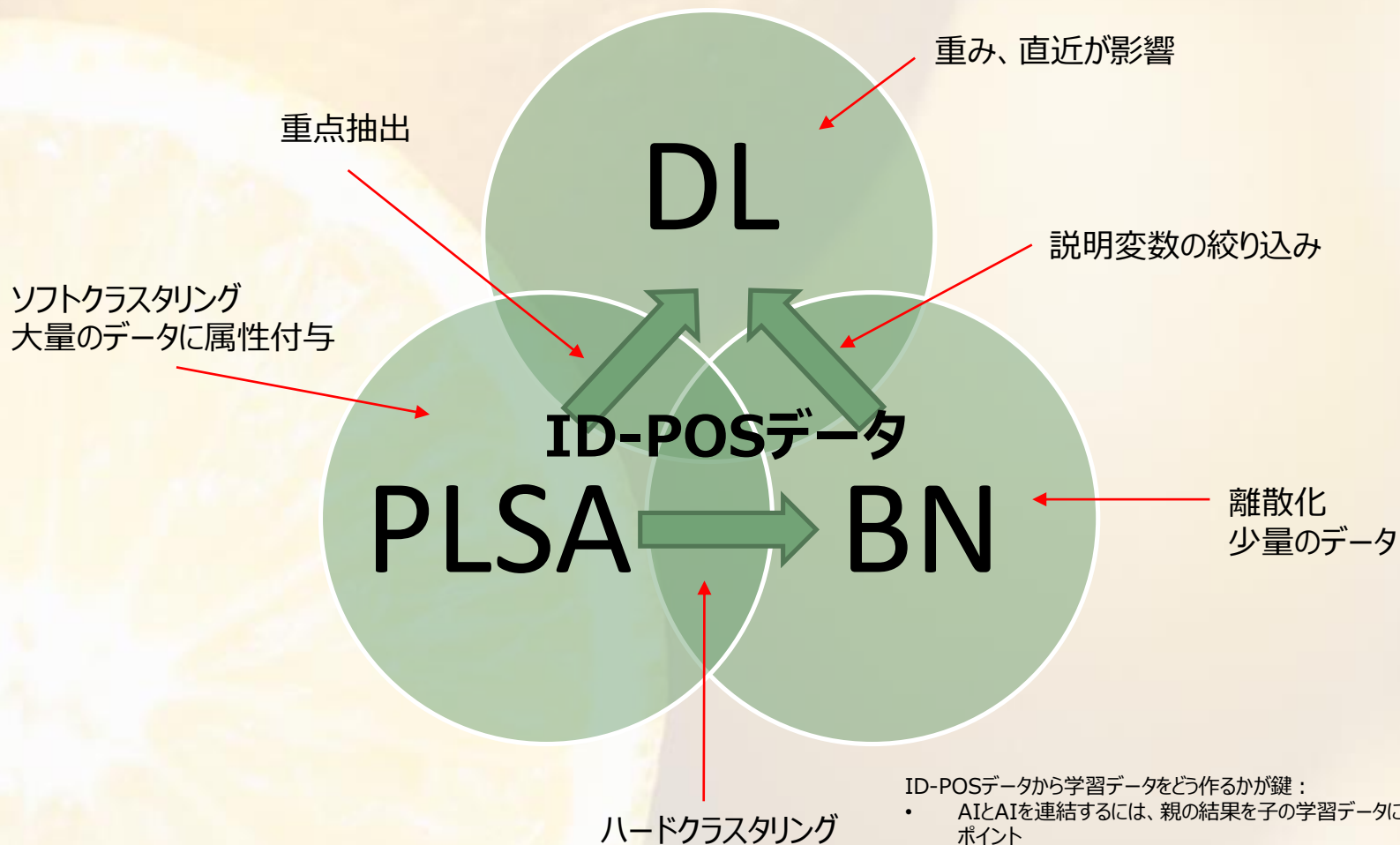


商品vs顧客のPLSA：

- 商品DNA vs 顧客DNA
 - 商品名に商品DNAが付加される
 - 商品のAI仮想クラスター
 - 顧客番号に顧客DNAが付加される
 - 顧客のAI仮想クラスター


商品分類（大中小、…）、顧客属性に
PLSA, BN等でのDNAをどう付与し、ID-POS分析に活用するか？

3. AIのハイブリッド化の課題



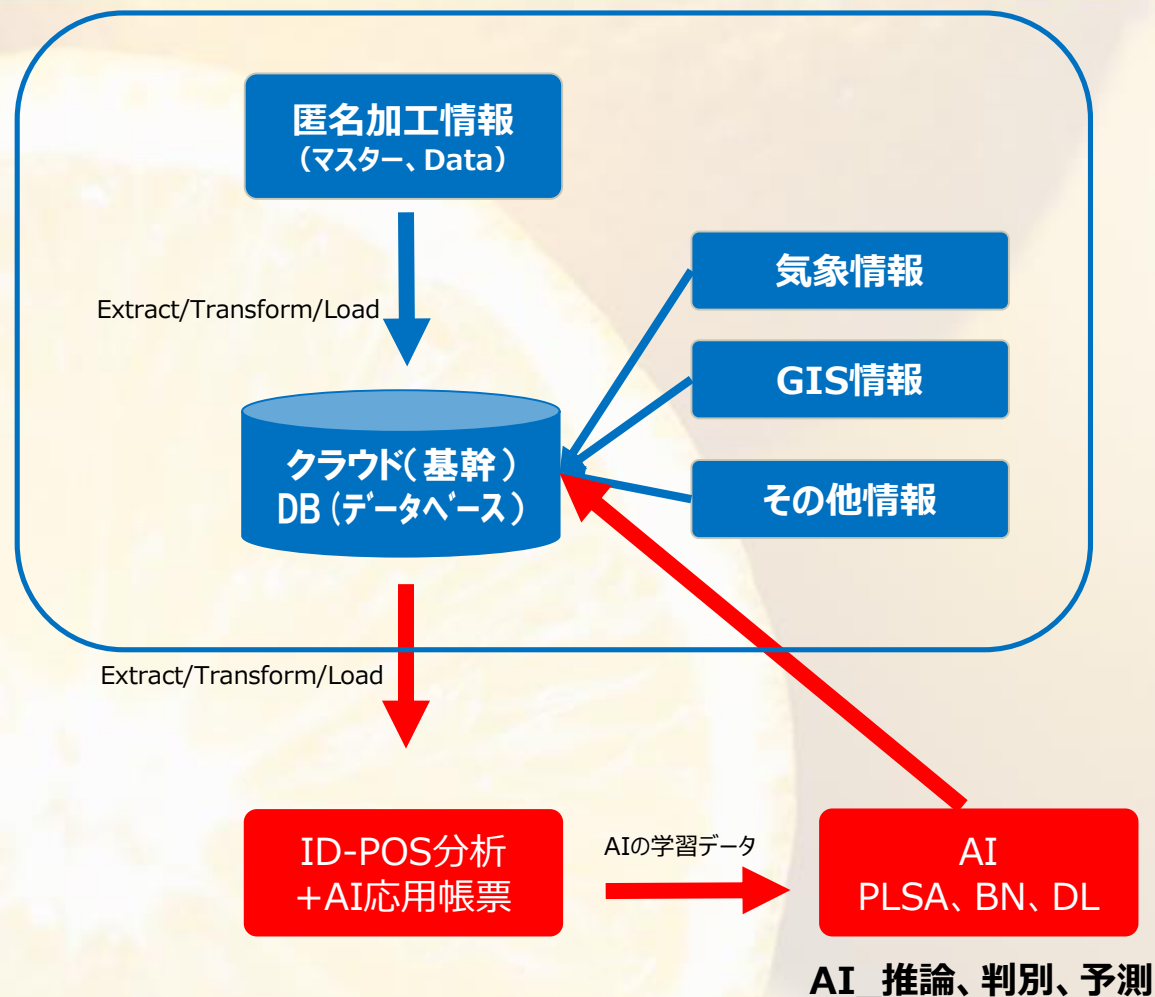
ID-POSデータから学習データをどう作るかが鍵：

- AIとAIを連結するには、親の結果を子の学習データに適合させることがポイント
 - 学習データの説明変数（条件）により精度の落差が激しい。
- 学習データの共通化が課題
 - そのままのデータを使用するか、分散化するか？



ID-POS分析は **AI**で進化する

1. 基幹システムとの連携、with AIがポイント



補足：

1. 匿名加工情報+α
 - 内閣府の個人情報保護委員会のガイドラインに沿って加工
 - 気象情報、GIS情報の提供企業選定
 - 人流データ、SNS等は今後の課題
2. クラウド&データベース
 - クラウド&データベース：
3. ID-POS分析システム
 - AIとの相互連携システムを開発
 - 学習データ作成システムを開発
4. AI
 - Withの段階では現有AIを活用
 - Inは別途開発か？
5. AI応用帳票
 - クラウドからSQLでAI&ID-POS分析に入れ、ここでAI関連の様々な帳票を制作

2.スパコンとの連携、ID-POS分析をABCIでいかに実践活用するか

AI橋渡しクラウド（AI Bridging Cloud Infrastructure、ABCI）

発表・掲載日：2018/06/26

ツイート いいね! 152

大規模AIクラウド計算システム「ABCI」がスパコン性能ランキング世!

—大規模で省電力のクラウド型計算システムで高度な人工知能処理を可能に—

ポイント

- 2018年8月1日運用開始：
- 1. 経産省「人工知能に関するグローバル研究拠点整備事業」（H28二次補正）の一環として整備
- 2. 我が国における産学官共同による**AI研究開発を加速するオープンイノベーションプラットフォーム**
- 3. 高い計算能力を活用した**人工知能技術の研究開発・実証社会実装の推進、AI分野の最重要課題への挑戦**が目的
- 4. **公的かつオープン**で、AI技術開発に特化した計算インフラとして**世界初の先進的システム**
- 5. **ディープラーニング**の学習速度の**世界記録**を大幅に更新
- 6. **約2,000人が利用**（2020年3月現在）

今後、公開されている予備演算モデルやオープンデータをABCI上で提供し、人工知能技術の研究開発用の大規模高速計算基盤として、産学官共同によるAI研究開発を加速し、AI分野の最重要課題に挑戦します。



左：ABCI計算ノードラックの一部、中央：冷却システムの一部、右：AIデータセンター棟

ソフトウェアスタック

ソフトウェア

※ABCIでは、HPC分野で定評のある最先端ミドルウェア、各種並列化コンパイラ、最新のGPU向け開発環境やライブラリ、各種デバッグフレームワークを提供します。

Operating System	CentOS RedHat Enterprise Linux
Job Scheduler	Univa Grid Engine
Container Engine	Docker Singularity
MPI	OpenMPI MVAPICH2-GDR MVAPICH2 Intel MPI
Development Tools	Intel Parallel Studio XE Cluster Edition PGI Professional Edition NVIDIA CUDA SDK GCC, Python, Ruby, R, Java, Scala, Lua, Perl
Deep Learning	Caffe, Caffe2, TensorFlow, Theano, Torch, PyTorch, CNTK, MXnet, Chainer, Keras NVIDIA GPU Cloud (NGC)
Big Data Processing	Hadoop, Spark

https://abci.ai/ja/about_abci/software.html