

実データを用いた業績判別を利用した企業の取引先選択に影響する特徴量 Effective Features for selecting business partners using Machine Learning with real data

向井 大誠*¹
Mukai Taisei

鳥山 正博*²
Toriyama Masahiro

寺野 隆雄*¹
Terano Takao

*¹ 東京工業大学
Tokyo institute of technology

*² 立命館大学
Ritsumeikan University

Recent rapid business market change requires that every firm should select or change their business partners to acquire resources. To analyze the effective features for selecting business partners, we study the real trade data using machine learning. The results suggest that effective features have partner's category information, growth indicators and partner's network indicators in high discriminant case.

1. はじめに

近年、商環境の変化が速い。速い商環境の変化に適応し収益を上げるために、企業は顧客のニーズを掴み、これに対応するためのリソースを提供してくれるパートナー(取引先)を迅速に選択する必要がある[Pathak 07][Wang 08][Mukai 12, 16]。

アプローチとして、企業が取引先を選択する際に、取引先のどのような要因が対象企業の直近の業績(年次)に効果があるかを分析する。取引のある企業間双方(取引ペア)の時系列な関連情報をユニークな特徴量として機械学習を行い、業績向上に効く特徴量の重要度を抽出した。対象企業から見た売り手、買い手の方向別に学習ができる点がポイントになる。高い判別力が学習できた実験では、取引先数や取引先の業種など会計的情報以外の要因が強く影響するケースが存在した。

2. 関連研究

企業のパートナー選択に関する研究は、パートナーを探索するもの、パートナーをあらかじめ与えられたパートナー候補から選ぶもの、既知のパートナー情報を元に学習・推薦するものがある。

パートナーを探索する研究は、企業のサービス流のパス[Chou 10][Niu 2011]を求めるもの、動的な企業間取引構造を扱うもの[Mukai 2016]、イノベーション[Araújo 09]に繋げる研究がある。これらは本研究のような実データに基づく学習・推薦を行っていない。

パートナーを候補から選ぶ研究[貝原 04][Wang 08][Behdani 13]は、基本サプライチェーンの研究の延長にある。パートナー候補は、同じ業種内で与えられ、選択に考慮する要素はすべて事前に与えられることから、よく知られた業種向けの研究と言える。本研究は、速い商環境の変化に適応するべく、パートナーの選択に必要な要素を学習し変化させる。

パートナーの情報を学習し推薦する研究は以下のようなものがある[You 08][Guo 09][Mori 12][伊藤 15]。前2者は QCT などサプライチェーンに特有の情報で機械学習しているが取引ペアの情報を扱っていない。後者は取引ペアの情報で正解を作っており当研究に近い。いずれのケースも特徴量に取引情報や WEB 情報を使用しており特徴があるが、取引リンクを予測する学習になっており、既存の取引ペアを正解にしているため、本研究のような環境の変化に迅速に適応することを強く意識した推薦ではない。

本研究は、実データを時系列に利用するため、商環境の変化の方向を反映できる。また、直近の新規取引ペアで学習するため迅速なパートナー選択に適用できる。

3. 手法

実際の新規取引ペアの情報を使用して、対象企業と取引先の情報がその後、対象企業の業績に有効に働いたかどうかを学習し、取引先のどのような情報(特徴量)が対象企業の業績向上に寄与したかを分析する。学習を新規取引に限定するため、既存の取引の影響を除外できる。分析手法は、特徴量の組み合わせを学習に反映させたいため機械学習の一つであるランダムフォレスト[Breiman 01]を使用する。特徴量の重要度を業績に効いた要因として検討する。

3.1 目的変数

新規取引のあった翌年の業績(plofit3)の前年(plofit2)との比較で目的変数を設定する。業績は税引き後利益を使用した。本来は経営者による操作が困難な売上か売上総利益を利用したかったが、サンプルデータが少なく計算困難なためこのようにした。税引き後利益は市場変動の影響を除外するために業界平均(Category ave)からの相対値を計算した。この値が前年より大きければ正例 1、小さければ負例 0 とした。

- $(\text{Profit3} - \text{Category ave3}) > (\text{Profit2} - \text{Category ave2}) \dots 1$
- $(\text{Profit3} - \text{Category ave3}) < (\text{Profit2} - \text{Category ave2}) \dots 0$

3.2 説明変数

本研究のユニークな点はここである。対象企業と取引先企業のペアを同時に考慮した変数選択になる。対象企業は自身が売り手のケースと買い手のケースになる場合を分けて分析するため、自身が売り手、買い手の立場で取引先の有効な要因(特徴)を検討できる(図 1)。また、それぞれ1年前の企業情報も用いるため、時間的な変化も考慮した学習になる。説明変数は[経済産業省 05]を参考にして、企業価値の計算に使用する変数を選択、表1の要素を両企業の情報として選択した。(尚、計算元の変数(売上(sales)、利益(profit)、労働者数(workers)、資本(capital))も説明変数として採用する。右横の数字は、2 が新規取引のあった年、1 がその前年の値に相当する。3 は正解に使用する年である。取引数は新規取引の年だけカウントした。)

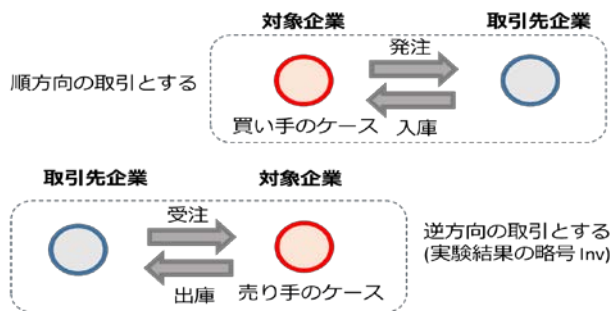


図 1: 学習データの企業ペアの関係 (対象企業が買い手と売り手のケースがあり、取引の1年前の情報も説明変数に使用する)

表 1: 企業価値計算を起因とした説明変数

収益性	profit2/sales2, profit1/sales1
成長性	(profit2-profit1)/profit1, (sales2-sales1)/sales1
効率性	sales2/capital2, sales1/capital1
安定性	workers2, workers1, outDout, outDin (売り手の出庫先数、入庫先数) inDout, inDin (買い手の出庫先数、入庫先数)
資本	capital2, capital1

*成長性、取引先の出庫・入庫数以外は前年の値も採用

また、これに加えて、対象企業と取引先の業種名をカテゴリカル変数として加える。例えば、自動車小売と自動車卸の取引であれば、自動車小売と自動車卸に 1 が立つ。同業種の取引の場合 2 が立つ。これは特定の取引業種の取引が業績に効くのか調べるためである。まとめると、以下の内容を説明変数として両企業の情報を採用する。

- ・新規取引のあった対象企業と取引先企業の情報(表 1)
- ・両企業の1年前の企業情報(取引情報はないケース)
- ・両企業の出庫先数、入庫先数
- ・両企業の業種(カテゴリカル変数、業種に 1 が立つ。同業種の取引の場合 2 が立つ)

3.3 学習サンプルと学習

学習サンプルは対象企業とその取引先の新規取引のペアをベースに作成するが、1社の取引先が多数になるケースがあるため、1社あたり利益の上位 10 件で打ち切る。また、合併の影響が業績に大きく影響している可能性が高いため、学習期間中に合併があったケースはすべて取り除いた。さらに、新規取引で絞る上に、業界平均を考慮するためサンプルが正例、負例のバランスの偏った不均衡データとなっている。そこで正解の少ない方に重みをかけた学習を行う。例えば、正例と負例の比が $a:b(a>b)$ であれば、 b に a/b 倍の重みをかける。なおランダムフォレストとテストのパラメータは以下に設定する。

表 2: ランダムフォレストとテストのパラメータ

木の数、木の深さ	500, 3(過学習を考慮)
最大特徴量の数	$d^{(1/2)}$... d は全特徴量数
情報量	ジニ係数
交差検証	$K=5$, leave-one-out 正答率 (oob 誤り率ではない), F 値

*交差検証はサンプルをシャフル後に Stratified K-Fold Cross Validation. を実行。機械学習は Scikit-learn in python3 を使用。

4. 実験

3 章で述べた目的変数と説明変数でランダムフォレストによる学習を行い、特徴量の重要度を取引先の選定に効く要因として検討する。対象企業の業種を中分類の”自動車自動車小売”にしばり学習する。これは学習サンプルがある程度あるためと、結果の考察が認識しやすいためである。まず、以下の実験条件で正答率、F 値を比較する。(尚、逆方向取引(対象企業が受注する取引)は、Inv の記号を打っている。順方向の取引(対象企業が発注する取引)は特に何も記載はない(図 1)。)

- ・説明変数に操作なし(略号 Base)
- ・正解を業界平均からの相対値で計算(略号 ave)
- ・合併データの削除(略号 ma)
- ・少ない解の方に負荷を掛けて学習(略号 W)

負荷を掛けて学習した結果の特徴量に着目する。また、学習の効果を見るために外挿テストを行う。加えて、正解に業界平均を考慮したことの効果、および合併データの削除の効果も考察する。

4.1 データ

株式会社帝国データバンク提供の信用調査報告書の取引先欄から構築した取引データと COSMOS2 を企業 id で紐付けして使用している。新規取引のあった年と前年の企業情報で学習し、その翌年の企業情報から正解ラベルを作成する(2014-2015(正解 2016))。また、業種カテゴリは小分類を使用し約 600 数十のカテゴリを使用する(年により違う)。外挿テストに関しては、2012-2013 を学習に使用して、2013-2014(正解 2015)および 2014-2015(正解 2016)の学習サンプルとラベルに対してテストしている。

5. 実験結果と考察

学習器の評価を図 2、外挿テストを図 3、4、特徴量の重要度を図 5-8 に示す。

5.1 実験結果

図 2 より、解の少ない方に負過学習の操作を行ったケースが順方向、逆方向ともに高い正答率と F 値を出している。取引先のカテゴリや成長性指標が強く効いていることがわかる。また、すべての操作を行ったケース(BaseAveMaW)の外挿テストでは、負荷学習を行ったケースのすべての指標が2年後の時点で下落している(図 3)。

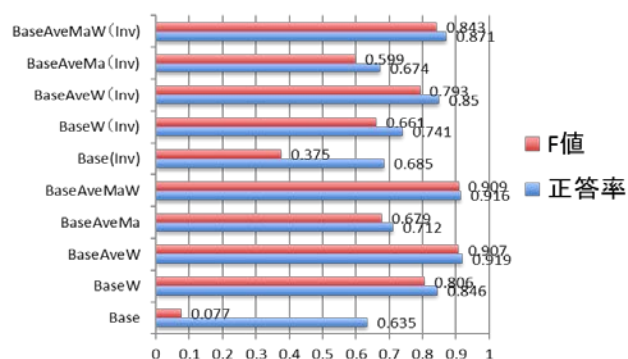


図 2: ランダムフォレストの交差検証結果(base が最も F 値が悪い、特に解の少ないほうに負荷をかけた学習時の際に高い判別力がある)

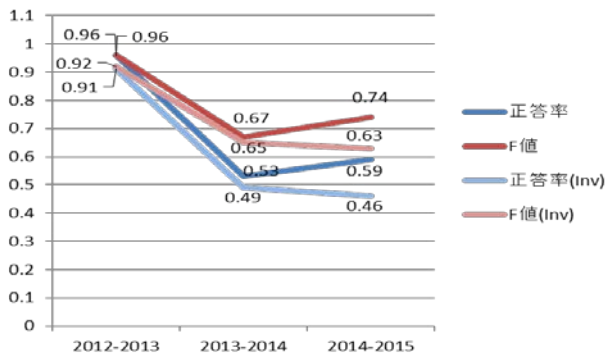


図 3:外挿テスト(負荷ありのケース、学習の翌年のサンプルでの予測時に成績が大きく下落している。)

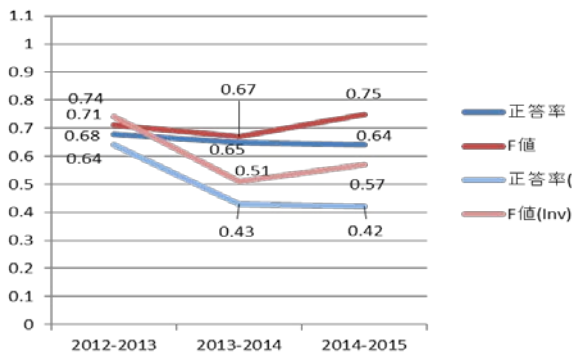


図 4:外挿テスト(負荷なしのケース、逆方向の取引のケースが下落している。負荷有のケースより下落幅が小さい)

同条件で負荷学習を行わなかったケースでは正答率、F値がそこまで下落していない(図 4)。また、負荷学習をしないと特徴量に業種カテゴリが強く効いていない(図 6)。業界平均を考慮したケースでは、同業種カテゴリが(BaseW:図 8)に比べてやや効いている(図 7)。他、全体を通して直近の売上、収益性指標および成長性指標は、対象企業、取引企業共に良く効いている。

業種カテゴリ				サンプル				201401-201501(正解201601)			
順方向取引:対象企業が買い手(対象企業の発注)の取引				逆方向取引:対象企業が売り手(対象企業の受注)の取引				正解			
正解	利益	業界平均	正負例比	正解	利益	業界平均	正負例比	正解	利益	業界平均	正負例比
正答率	0.916	合併	削除	正答率	0.919	合併	存在	正答率	0.919	合併	存在
F値	0.909	負荷	あり	F値	0.907	負荷	あり	F値	0.907	負荷	あり
no	特徴量名	重要度		no	特徴量名	重要度		no	特徴量名	重要度	
0	4710.自動車小売	0.1098		0	4710.自動車小売	0.1055		0	4710.自動車小売	0.1277	
1	(sales2-sales1)/sales1	0.0812		1	(sales2-sales1)/sales1_inv	0.0667		1	(sales2-sales1)/sales1_inv	0.0606	
2	sales2	0.0598		2	(profit2-profit1)/profit1	0.0562		2	(profit2-profit1)/profit1_inv	0.0604	
3	(sales2-sales1)/sales1_inv	0.0587		3	inDin	0.0540		3	(profit2-profit1)/profit1	0.0582	
4	(profit2-profit1)/profit1	0.0585		4	(sales2-sales1)/sales1	0.0526		4	(sales2-sales1)/sales1	0.0563	
5	4072.自動車部品付属品卸	0.0514		5	(profit2-profit1)/profit1_inv	0.0484		5	outDout	0.0442	
6	outDout	0.0457		6	salesInv2	0.0417		6	inDin	0.0433	
7	profit2/sales2	0.0452		7	profit1/sales1_inv	0.0390		7	salesInv2	0.0420	
8	profit	0.0363		8	salesInv	0.0385		8	profit2/sales2_inv	0.0360	
9	(profit2-profit1)/profit1_inv	0.0362		9	profit2/sales2_inv	0.0350		9	profit1/sales1_inv	0.0337	
10	profit2	0.0317		10	profit	0.0326		10	profit2	0.0324	
11	sales	0.0283		11	profit2	0.0306		11	salesInv	0.0323	
12	inDin	0.0267		12	inDout	0.0290		12	profit	0.0289	
13	capital	0.0229		13	profitInv2	0.0285		13	outDout	0.0269	
14	workers	0.0223		14	workersInv2	0.0270		14	workersInv	0.0250	
15	capital2	0.0202		15	profitInv	0.0247		15	6310.一般貨物自動車運送	0.0220	
16	4071.自動車卸	0.0199		16	capitalInv2	0.0218		16	profitInv2	0.0204	
17	workers2	0.0196		17	workersInv	0.0218		17	workersInv2	0.0200	
18	4220.代理商・仲立業	0.0159		18	outDout	0.0186		18	capitalInv2	0.0192	
19	profitInv	0.0151		19	6310.一般貨物自動車運送	0.0167		19	profitInv	0.0180	
20	4720.各種二輪車小売	0.0138		20	sales	0.0155		20	capital	0.0174	
21	inDout	0.0133		21	capital	0.0152		21	capitalInv	0.0160	
22	profitInv2	0.0115		22	capital2	0.0146		22	capital2	0.0142	
23	5520.損害保険	0.0113		23	capitalInv	0.0139		23	7440.自動車賃貸	0.0127	
24	outDin	0.0113		24	outDin	0.0123		24	outDin	0.0108	
25	workersInv2	0.0097		25	workers	0.0121		25	workers	0.0108	
26	profit1/sales1_inv	0.0091		26	7440.自動車賃貸	0.0102		26	workers2	0.0101	
27	profit1/sales1	0.0090		27	4072.自動車部品付属品卸	0.0101		27	4072.自動車部品付属品卸	0.0095	
28	workersInv	0.0081		28	sales2	0.0096		28	profit2/sales2	0.0084	
29	8523.情報提供サービス	0.0080		29	workers2	0.0094		29	8210.自動車整備	0.0084	
30	capitalInv	0.0079		30	profit1/sales1	0.0089		30	sales	0.0079	
31	profit2/sales2_inv	0.0072		31	profit2/sales2	0.0069		31	sales2	0.0077	
32	capitalInv2	0.0063		32	8210.自動車整備	0.0068		32	profit1/sales1	0.0073	
33	salesInv2	0.0055		33	4220.代理商・仲立業	0.0063		33	4220.代理商・仲立業	0.0057	
34	salesInv	0.0050		34	1770.管工事業	0.0037		34	8130.有線放送	0.0048	
35	1610.一般土木建築工事業	0.0044		35	6212.タクシーハイヤー業	0.0037		35	1620.土木工事業	0.0045	

図 5:特徴量の重要度(base+ave+ma+W、すべての操作を行ったケース。成長性指標や業種カテゴリが効いている。赤:数値データより上に来た業種カテゴリ(上位 35 以上)、緑:取引数;outDout 発注側の受注社数、outDin 発注側の発注社数、先頭 in は受注側で同様)

業種カテゴリ				サンプル				201401-201501(正解201601)			
順方向取引:対象企業が買い手(対象企業の発注)の取引				逆方向取引:対象企業が売り手(対象企業の受注)の取引				正解			
正解	利益	業界平均	正負例比	正解	利益	業界平均	正負例比	正解	利益	業界平均	正負例比
正答率	0.712	合併	削除	正答率	0.674	合併	削除	正答率	0.674	合併	削除
F値	0.679	負荷	なし	F値	0.599	負荷	なし	F値	0.599	負荷	なし
no	特徴量名	重要度		no	特徴量名	重要度		no	特徴量名	重要度	
0	sales2	0.0995		0	profit2/sales2_inv	0.0886		0	profit2/sales2_inv	0.0886	
1	profit2/sales2	0.0760		1	salesInv2	0.0697		1	salesInv2	0.0697	
2	profit	0.0733		2	salesInv	0.0628		2	salesInv	0.0628	
3	profit2	0.0619		3	workersInv	0.0512		3	workersInv	0.0512	
4	(sales2-sales1)/sales1	0.0551		4	profit1/sales1_inv	0.0504		4	profit1/sales1_inv	0.0504	
5	sales	0.0521		5	workersInv2	0.0466		5	workersInv2	0.0466	
6	workers	0.0521		6	capitalInv2	0.0462		6	capitalInv2	0.0462	
7	capital2	0.0475		7	profitInv	0.0420		7	profitInv	0.0420	
8	workers2	0.0474		8	capitalInv	0.0411		8	capitalInv	0.0411	
9	capital	0.0415		9	profitInv2	0.0349		9	profitInv2	0.0349	
10	profit1/sales1	0.0310		10	(sales2-sales1)/sales1_inv	0.0298		10	(sales2-sales1)/sales1_inv	0.0298	
11	capitalInv2	0.0216		11	(profit2-profit1)/profit1_inv	0.0291		11	(profit2-profit1)/profit1_inv	0.0291	
12	capitalInv	0.0195		12	inDout	0.0284		12	inDout	0.0284	
13	workersInv	0.0178		13	sales	0.0244		13	sales	0.0244	
14	profit1/sales1_inv	0.0178		14	capital	0.0235		14	capital	0.0235	
15	salesInv2	0.0168		15	workers	0.0230		15	workers	0.0230	
16	outDout	0.0164		16	inDin	0.0215		16	inDin	0.0215	
17	workersInv2	0.0162		17	sales2	0.0210		17	sales2	0.0210	
18	profit2/sales2_inv	0.0156		18	capital2	0.0196		18	capital2	0.0196	
19	profitInv	0.0152		19	profit2	0.0191		19	profit2	0.0191	
20	inDin	0.0150		20	profit	0.0181		20	profit	0.0181	
21	outDin	0.0137		21	workers2	0.0156		21	workers2	0.0156	
22	4074.輸送用機械器具卸	0.0133		22	profit2/sales2	0.0148		22	profit2/sales2	0.0148	
23	salesInv	0.0128		23	profit1/sales1	0.0141		23	profit1/sales1	0.0141	
24	inDout	0.0107		24	outDout	0.0130		24	outDout	0.0130	
25	(sales2-sales1)/sales1	0.0101		25	4072.自動車部品付属品卸	0.0129		25	4072.自動車部品付属品卸	0.0129	
26	(profit2-profit1)/profit1_inv	0.0091		26	(sales2-sales1)/sales1	0.0128		26	(sales2-sales1)/sales1	0.0128	
27	4710.自動車小売	0.0087		27	6310.一般貨物自動車運送	0.0119		27	6310.一般貨物自動車運送	0.0119	
28	profitInv2	0.0080		28	outDin	0.0058		28	outDin	0.0058	
29	(profit2-profit1)/profit1	0.0065		29	8899.他の事業サービス	0.0053		29	8899.他の事業サービス	0.0053	
30	4220.代理商・仲立業	0.0051		30	1770.管工事業	0.0050		30	1770.管工事業	0.0050	
31	3711.自動車製造	0.0040		31	(profit2-profit1)/profit1	0.0049		31	(profit2-profit1)/profit1	0.0049	
32	4062.産業用電気機器卸	0.0037		32	4710.自動車小売	0.0049		32	4710.自動車小売	0.0049	

図 6:特徴量の重要度(base+ave+ma、図 5 と比較して負過学習だけを行わなかったケース。業種カテゴリが数値情報ほど効いていない。図4と同じモデルの 2014-2015 のサンプルでの学習結果)

業種カテゴリ				サンプル				201401-201501(正解201601)			
販方向取引:対象企業が買い手(対象企業の発注)の取引				逆方向取引:対象企業が売り手(対象企業の受注)の取引							
正解	利益	業界平均	正負例比	385:451	正解	利益	業界平均	正負例比	196:258		
正答率	0.919	合併	存在		正答率	0.850	合併	存在			
F値	0.907	負荷	あり		F値	0.793	負荷	あり			

no	特徴量名	重要度
0	4710.自動車小売	0.1277
1	(sales2-sales1)/sales1	0.0829
2	(sales2-sales1)/sales1_inv	0.0561
3	(profit2-profit1)/profit1	0.0555
4	sales2	0.0496
5	outDout	0.0438
6	profit2/sales2	0.0423
7	(profit2-profit1)/profit1_inv	0.0409
8	4072.自動車部品付属品卸	0.0384
9	profit	0.0375
10	profit2	0.0305
11	inDin	0.0259
12	sales	0.0221
13	capital	0.0206
14	outDin	0.0201
15	profitInv	0.0195
16	4220.代理商・仲立業	0.0192
17	profitInv2	0.0168
18	capital2	0.0162
19	5520.損害保険	0.0158
20	4071.自動車卸	0.0156
21	workers	0.0156
22	4720.各種二輪車小売	0.0144
23	inDout	0.0130
24	workers2	0.0124
25	workersInv	0.0114
26	profit1/sales1	0.0112
27	8253.情報提供サービス	0.0111
28	capitalInv2	0.0099
29	capitalInv	0.0098
30	salesInv2	0.0078
31	4074.輸送用機械器具卸	0.0071
32	salesInv	0.0071
33	profit2/sales2_inv	0.0069
34	workersInv2	0.0068
35	profit1/sales1_inv	0.0053
36	4073.中古自動車卸	0.0044

no	特徴量名	重要度
0	4710.自動車小売	0.1043
1	(sales2-sales1)/sales1_inv	0.0606
2	(profit2-profit1)/profit1_inv	0.0604
3	(profit2-profit1)/profit1	0.0582
4	(sales2-sales1)/sales1	0.0583
5	inDout	0.0442
6	inDin	0.0433
7	salesInv2	0.0420
8	profit2/sales2_inv	0.0360
9	profit1/sales1_inv	0.0337
10	profit2	0.0324
11	salesInv	0.0323
12	profit	0.0289
13	outDout	0.0269
14	workersInv	0.0250
15	6310.一般貨物自動車運送	0.0220
16	profitInv2	0.0204
17	workersInv2	0.0200
18	capitalInv2	0.0192
19	profitInv	0.0180
20	capital	0.0174
21	capitalInv	0.0160
22	capital2	0.0142
23	7440.自動車賃貸	0.0127
24	outDin	0.0108
25	workers	0.0108
26	workers2	0.0101
27	4072.自動車部品付属品卸	0.0095
28	profit2/sales2	0.0084
29	8210.自動車整備	0.0084
30	sales	0.0079
31	sales2	0.0077
32	profit1/sales1	0.0073
33	4220.代理商・仲立業	0.0067
34	8130.有線放送	0.0048
35	1620.土木工事業	0.0045
36	4720.各種二輪車小売	0.0042

業種カテゴリ 47 自動車自転車小売				サンプル 201401-201501(正解201601)			
販方買取対象企業が買い手(対象企業の発注)の取引				法方買取対象企業が売り手(対象企業の受注)の取引			
正解	利益	正負割合	316.520	正解	利益	正負割合	175.279
正答率	0.846	合併	存在	正答率	0.808	合併	存在
F値	0.741	負荷	あり	F値	0.661	負荷	あり

no	特徴量名	重要度	no	特徴量名	重要度
0	4710 自動車小売	0.1234	0	4710 自動車小売	0.0999
1	(sales2-sales1)/sales1	0.0638	1	(sales2-sales1)/sales1_inv	0.0699
2	(profit2-profit1)/profit1	0.0599	2	(profit2-profit1)/profit1	0.0635
3	4072_自動車部品付属品卸	0.0575	3	(sales2-sales1)/sales1	0.0573
4	(sales2-sales1)/sales1_inv	0.0514	4	inDin	0.0570
5	(profit2-profit1)/profit1_inv	0.0463	5	(profit2-profit1)/profit1_inv	0.0538
6	outDout	0.0435	6	inDout	0.0521
7	profit	0.0397	7	profit	0.0441
8	sales2	0.0382	8	profit2	0.0349
9	profit2/sales2_inv	0.0299	9	salesInv2	0.0314
10	profit2	0.0289	10	profit2/sales2_inv	0.0295
11	profit2/sales2	0.0288	11	profit1/sales1_inv	0.0286
12	outDin	0.0282	12	outDout	0.0280
13	profit1/sales1_inv	0.0230	13	6310_一般貨物自動車運送	0.0279
14	inDin	0.0225	14	salesInv	0.0195
15	4220_代運商・仲立業	0.0197	15	7440 自動車賃貸	0.0188
16	inDout	0.0192	16	capital	0.0182
17	profitInv	0.0173	17	profitInv	0.0162
18	workers2	0.0162	18	outDin	0.0155
19	profit1/sales1	0.0157	19	workers	0.0151
20	4071_自動車卸	0.0151	20	workersInv2	0.0141
21	workersInv2	0.0145	21	profit2/sales2	0.0135
22	4720_各種二輪車小売	0.0143	22	4072_自動車部品付属品卸	0.0134
23	5520_損害保険	0.0131	23	profit1/sales1	0.0133
24	workersInv	0.0127	24	capitalInv2	0.0121
25	profitInv2	0.0125	25	workers2	0.0117
26	sales	0.0124	26	profitInv2	0.0109
27	workers	0.0118	27	capital2	0.0109
28	0523_情報提供サービス	0.0108	28	sales	0.0103
29	salesInv2	0.0107	29	capitalInv	0.0100
30	capitalInv2	0.0101	30	workersInv	0.0097
31	salesInv	0.0094	31	sales2	0.0083
32	capitalInv	0.0089	32	8210 自動車整備	0.0081
33	capital2	0.0082	33	1620 土木工事業	0.0054
34	capital	0.0078	34	8699_他の事業サービス	0.0052
35	4074_輸送用機械器具卸	0.0050	35	5520_損害保険	0.0050

図 8:特徴量の重要度(baseW、操作なしのケース、売り手方向の取引に取引数が強く効いている。また、買い手方向の取引に同業でない業種カテゴリが効いているものがある。)

また、業界平均を考慮して学習すると同業種カテゴリが上がってくることから、同業種間の取引が業績向上に影響していることがわかる。さらに合併を省くことにより取引数の重要度の変動を確認できるが、これは取引数の急増を省く形になるので直感に反しない(図5)。順方向取引に関しては、卸や保険・情報サービスへの発注が効き、逆方向取引に関しては車両を利用するサービスからの受注が効いている。いずれも直感に反しないが、他の取引よりも重要であると一意に見抜ける取引ではない。

6. まとめ

解の少ないように負荷をかけた学習をすると、対象の業種に過学習することがわかった。しかし、学習した翌年の業績に効く特徴量が抽出できた。過学習であっても早い環境適応に効く要因を知るという点で意味のある過学習であると考察する。また、特に学習時に負荷をかけなければ、年次で判別力が維持しそうな気配であるため、これら特徴量の重要度を考慮して説明変数を取捨すると長期的に有効かもしれない。但し、学習データが多くのケースで不均衡データになるため、判別力は業種カテゴリごとにサンプルの不均衡を割り引いて考察する必要がある。

今後の展開として、正解を連続値や複数の離散値にしてレベル分けをする学習をし、取引先を厳格に判別し選択に生かす筋が考えられる。また、カテゴリ別の比較を行い、業界をまたいだ取引先推薦の可能性を検討したい。

謝辞

本研究の一部は、東京工業大学帝国データバンク先端データ解析共同研究講座研究講座の支援を受けた。感謝の意を表する

参考文献

- [Araújo 09] Araújo, T., & Mendes, R. V.: Innovation and self-organization in a multi-agent model, *Advances in Complex Systems*, vol. 12, no. 02, pp. 233-253, (2009).
- [Breiman 01] Breiman, L.: Random forests. *Machine learning*, vol 45, no 1, pp. 5-32, (2001).
- [Behdani 13] Behdani, B., van Dam, K. H., and Lukszo, Z.: Agent-Based Models of Supply Chains, In *Agent-Based Modelling of Socio-Technical Systems*, pp. 151-180, Springer Netherlands, (2013).
- [Chou 10] M. Chou and G. Chua: Design for process flexibility, Efficiency of the long chain and sparse structure, *Operations research*, vol. 58, no. 1, pp. 43-58, (2010).
- [Guo 09] Guo, X., Yuan, Z., & Tian, B.: Supplier selection based on hierarchical potential support vector machine. *Expert Systems with Applications*, vol 36, no. 3, pp. 6978-6985, (2009).
- [伊藤 15] 伊藤諒, 森純一郎.: テキスト情報を用いた企業間取引推薦システムの構築と評価 (言語理解とコミュニケーション)- (第7回テキストマイニング・シンポジウム). 電子情報通信学会技術研究報告= IEICE technical report: 信学技報, vol. 115, no. 222, pp. 43-46, (2015).
- [貝原 04] 貝原 俊也 and 藤井進: サプライチェーン環境下におけるバーチャルエンタプライズ型ビジネスモデルの特性評価, *日本機械学会論文集(C編)*, vol. 70, no. 690, pp. 539-546, (2004).
- [経済産業省 05] 経済産業省・法務省.: 企業価値・株主共同の利益の確保又は向上のための買収防衛策に関する指針, *経済産業省*, (2005).
- [Mori 2012] J. Mori, Y. Kajikawa, H. Kashima, and I. Sakata,:Machine learning approach for finding business partners and building reciprocal relationships, *Expert Systems with Applications*, vol.32, pp.10402-10407, (2012).
- [向井 12] 向井大誠, 諏訪博彦, 太田敏澄.: 主体性のあるリソースで構成する環境適応組織の仕組み, *社会情報学会 (SSI) 学会大会研究発表論文集*, pp.209-214, (2012).
- [Mukai 16] Mukai,T., Terano,T.: Modeling decentralized inter-organizational business structures through agent-based simulation. In *World Automation Congress (WAC)*, pp. 1-8, IEEE, (2016).
- [Niu 2011] Niu, H., Guo, P., Lu, Y., & Chen, H.: Efficient SOA Business Process Reconstruction for Disaster Recovery. In *Internet Technology and Applications (iTAP)*, International Conference, pp. 1-4, IEEE, (2011).
- [Pathak 2007] Pathak, S. D., Day, J. M., Nair, A., Sawaya, W. J., & Kristal, M. M.: Complexity and adaptivity in supply networks: building supply network theory using a complex adaptive systems perspective. *Decision Sciences*, vol. 38, no. 4, pp. 547-580, (2007).
- [Wang 08] Wang, M., Liu, J., Wang, H., Cheung, W. K., and Xie, X.: On-demand e-supply chain integration: A multi-agent constraint-based approach, *Expert Syst. Appl.*, vol. 34, no. 4, (2008).
- [You 08] You, Z., Si, Y. W., Zhang, D., Zeng, X., Leung, S. C., Li, T. : A decision-making framework for precision marketing. *Expert Systems with Applications*, vol.42, no.7, pp. 3357-3367, (2015).