

科学研究費助成事業 研究成果報告書

平成 29 年 5 月 17 日現在

機関番号：24403

研究種目：基盤研究(C) (一般)

研究期間：2013～2016

課題番号：25350343

研究課題名(和文) フィードバックプロセスを採用した実践的データマイニング教育プログラムの開発と実践

研究課題名(英文) Education Program for Data Mining incorporating feedback process

研究代表者

森田 裕之 (MORITA, HIROYUKI)

大阪府立大学・人間社会システム科学研究科・教授

研究者番号：80295732

交付決定額(研究期間全体)：(直接経費) 3,900,000円

研究成果の概要(和文)：本研究では、データマイニングを活用した実践的な問題解決型教育を効果的に実施するため、実験店舗におけるPOSデータを分析して店舗実験を実施し、そこから得られるデータを学生にフィードバックして活用することで、学生が実践的な状況観察能力、データ分析能力、そして総合的な問題解決能力を習得可能な教育プログラムを開発することが目的である。

研究期間において、上記プログラムを開発・実施すると共に、それを評価するためのアンケート票を設計・実施し、学生の能力向上について確認を行った。

研究成果の概要(英文)：As many people require business applications for Big-data, human resources who can perform data-mining enough, are needed. In order to develop such human resources, we have to provide practical data-mining education program. In the educational program, we should teach a lot of knowledge. They include several data-mining algorithms and another knowledge such as marketing. And we should handle total data-mining processes using such knowledge. This study proposes an educational program for practical data-mining. In the program, a method to evaluate capabilities of students, is proposed, and data which are collected by the method are analyzed.

研究分野：経営情報科学

キーワード：データマイニング 教育

1. 研究開始当初の背景

ICTの発達に伴い、従来、蓄積及び処理することが困難であったビッグデータを、安価かつ容易に利用可能となってきた。そのため実際にデータをマイニングして有効なビジネスソリューションを提案することが可能な人材が求められている。これには、単に分析手法を教育するだけでは十分とは言えない。その理由は、次のような点があげられる。a) 効果的なデータマイニングを実践するためには、問題を発掘し、必要なデータを収集するプロセスから、分析を踏まえて結果を解釈し、実践方法を検討するまでの一連の問題解決プロセスを理解し、その中で必要な分析を実行しなければならない。b) 実際のデータでは、欠損値はもちろん、データ形式など、手法を適用するまでのデータクリーニングおよび変数の設計の作業が不可欠である。そのため、実際のデータの利用経験が必要不可欠となる。以上の理由から、効果的にビジネス分野におけるデータマイニングを教育するためには、1) 経営戦略論やマーケティングなどの経営学的内容の理解と、統計学や情報工学的な分析手法の理解を複合的に行う必要があり、2) 実際のデータを用いて、問題発掘から変数の設計、データ分析、結果の解釈、そして得られた知見からのビジネスアクションの考察が必要不可欠である。また、3) 実際のビジネスアクションから得られる結果を更にデータとして収集し、その効果を把握すると共に、よりよいアクションを実践できるようなフィードバックプロセスを導入することが重要であると言える。

2. 研究の目的

本研究では、データマイニングを活用した実践的な問題解決型教育を効果的に実施するため、実験店舗におけるPOSデータとともに、各種の店舗実験の実施し、そこから得られるデータを学生にフィードバックして活用することで、学生が実践的な状況観察能力、データ分析能力、そして総合的な問題解決能力を習得可能な教育プログラムを開発することが目的である。

3. 研究の方法

実施年度中においては、店舗実験までを取り入れた実施教育プログラムを実践しながら、フィードバックプロセスの教育プログラムへの取り込み、及びその効果的な実施方法について検討を行い、これまでの教育プログラムのフローの見直しを行った。そして学生に対する教育効果の把握を行うためのアンケート票の設計変更についても作業を進めた。これまでどちらかと言えば、ビジネスソリューションの提案までのプロセスに時間を多く配分していたため、その点を明らかにするポイントに少しアンケート項目が集中していたが、実践およびそのフィードバック

部分に関しても、その効果を把握できるような項目を追加することによって、全体的な効果を確認できるように改善を行った。そしてその再設計を行った教育プログラムを前期と後期の2つの講義科目を利用して実施しながら、前年度に作成したアンケート票にもとづいて学生の能力取得を確認するデータの収集も実施した。またそのデータを分析することによって、教育プログラムの評価を実施した。

4. 研究成果

まず評価するための質問票としては、表1のようなアンケート項目に加えて、表2のような項目を整理して追加設計を行った。

表1 社会人基礎力に基づく質問項目

分類	番号	質問項目
ア ク シ ヨ ン カ	Q1	問題(課題)に直面した際、自分で進んで取り組み解決しようとしている。
	Q2	グループで行動する際、リーダーになることが多い。
	Q3	自分1人では解決できない問題に直面した際、他人を引き込んで協力を得ることができる。
	Q4	グループを統率することは得意だと思う。
	Q5	到達目標を予め低めに設定することが多い。
	Q6	設定した目的は、ほぼ確実に達成することができる。
シ ン キ ン グ カ	Q7	与えられた状況から、目標とのギャップを把握することは得意である。
	Q8	目標とのギャップを埋める方法を提案することは得意である。
	Q9	目標が与えられたら、それを達成するためのプロセスを見通すことができる。
	Q10	目標を達成するための計画を綿密に設計するほうだ。
	Q11	新たなアイデアや発明は得意である。
	Q12	問題を解決する際、人とは違う方法で解決しようと意識している。
チ ー ム ワ ー ク カ	Q13	自分の考えを人に伝えるのは得意である。
	Q14	相手の考えを理解するように努力している。
	Q15	自分とは異なる意見を持つものを許容し、より良い解を探索することができる。
	Q16	グループ行動するとき、グループ内の状況と、グループ外の状況の両方を把握することができる。
	Q17	約束したことは守るほうだ。
	Q18	ストレスがかかった際、うまく受け流すことができると思う。

これらの質問票を用いて、データマイニング教育を実施する際に、その前後において、同様の質問を実施して、教育の効果を分析した。内容としては、概ね、期待した分野の能力を学生が取得していることが確認され、教育プログラムとして実施することが適切であることが確認された。特に、情報処理能力、データ分析力、そしてデータマイニング能力に

については全体として大きな能力獲得を確認できたことは、大きな成果であったと思われる。これまでの研究期間において、教育プログラムの見直し、および取得能力を確認するためのアンケートの実施などを改善しながら行ってきた。

表2 データマイニング関係で追加した質問項目

分類	番号	質問項目
情報処理能力	Q19	エクセルを利用して、データを処理することができる。(データの並べ替え、フィルター、値の検索など)
	Q20	エクセルを利用して、データを集計することができる。
	Q21	エクセルを利用して、データを視覚化(グラフ作成)することができる。
	Q22	コマンド(コンピュータ言語でもシェルコマンドでもなんでも可)を使って、データ処理ができる。
データ解析力	Q23	コマンドを使ってデータを集計することができる。
	Q24	エクセル以外でデータを視覚化するツールが利用できる。
	Q25	何らかの数値を予測するモデルを作ることができる。
データマインニング力	Q26	何らかの所属クラスを予測するモデルを作ることができる。
	Q27	与えられたデータを、適切なグループに分けることができる。
	Q28	何らかの方法でデータ間の関係性を示すことができる。
	Q29	データから何らかの異常な状態を検知することができる。
データ分析力	Q30	データと目標が与えられたら、そこから目標達成するための課題を識別することができる。
	Q31	識別された課題を解決するために適切なデータセットを作成することができる。
	Q32	データセットを適切な分析手法を利用して計算し、結果を解釈することができる。
データマインニング力	Q33	解釈した結果から、識別した課題を解決するためのソリューションを導き出すことができる。
	Q34	ソリューションを実施計画レベルに変換し、ソリューションを実行できる。
データマインニング力	Q35	実行結果をフィードバックして、これまでのプロセスをリファインできる。

もう少し詳細に結果をみってみる。能力分類および全体に対して能力をどれくらい獲得しているかという割合を表す獲得率を以下のように定義する。

定義：獲得率 = (各能力分類 or 全体の合計値) / (各能力分類 or 全体の満点値)

この獲得率を前期と後期でまとめたものが表3であり、そのうち前期の獲得率の開始時と終了時の違いについてグラフ化したものが図1である。

表3 前期と後期の獲得率の変化

能力分類	前期		後期	
	開始時	終了時	開始時	終了時
アクション力	72.5%	70.5%	71.5%	74.3%
シンキング力	69.6%	72.1%	69.4%	69.4%
チームワーク力	75.4%	79.2%	84.7%	79.9%
情報処理能力	56.5%	73.7%	63.9%	71.5%
データ分析力	35.3%	59.2%	65.8%	65.0%
データマイニング力	37.1%	62.3%	55.6%	62.5%
全体	59.4%	71.5%	68.6%	70.6%

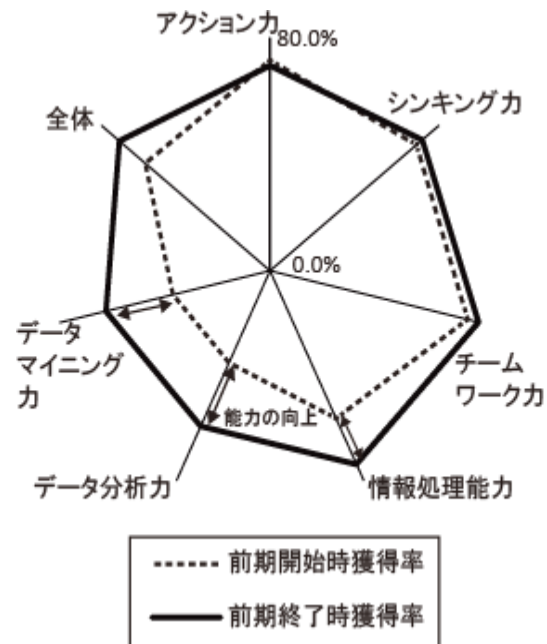


図1 前期における能力の向上

前期においては、“情報処理力”、“データ分析力”、および“データマイニング力”については、かなり大きな能力の向上が確認され、学生の主観的な能力獲得という点では、期待通りの結果が得られていることがわかる。

次に個人の各能力の平均値を求め、各期間で能力間の相関係数を計算して、関係性を考察する。図2は、能力間に有意な相関関係が確認される部分だけ枝を張って無向グラフで表現したものである。図より、“データ分析力”、“情報処理能力”、そして“データマイニング力”の間には複数の期間について強い関係が確認される。またこれらは、“シンキング力”とも関係が確認され、密接につながりがある。

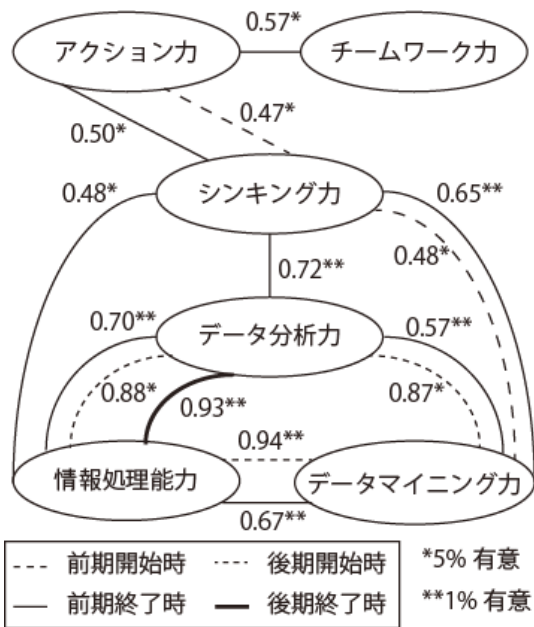


図2 主観的評価値間の相関関係

最後に主観的な能力向上の実感だけでなく、客観的な評価値の向上も確認するため回帰分析を行い、表4の結果を得た。これは前期のデータマイニングの講義において、終了時の値から開始時の値を差し引いた値である Q1~35 の差分を説明変数に、目的変数をレポートの点数(50点満点)として重回帰モデルを作成した結果である。

表8 レポートの成績と能力向上との関係

回帰統計	
重決定 R ²	0.38
補正 R ²	0.32
有意 F	0.008
観測数	23

	係数	t 値	P-値
切片	32.80	28.65	1.04E-17
差 Q15	2.60	2.31	0.032
差 Q20	3.28	2.48	0.022

モデルは 1%有意であり、Q15 と Q20 がプラスで有意な説明変数であるため、これらの値がプラスになった学生は、レポートの点数もよかったと言える。Q15 は“チームワーク力”の中の、“自分とは異なる意見を持つものを許容し、より良い解を探索することができる”という質問項目であり、Q20 は、“情報処理能力”の中の“データを集計することができる”という質問項目であった。

以上のように、店舗のデータを利用した実践的な教育を行い、上述のような質問項目を使用して、講義の前後にアンケートを実施す

れば、その教育成果を確認できることを示した。これらの結果より、当初予定していた目的を達成したものと考えることができる。

<引用文献>

森田裕之, "実践的なデータマイニング教育と学生の獲得能力に関する研究 (<特集> 情報教育の新展開)", 日本情報経営学会誌, 36, 2, pp.51-62, 2015, 日本情報経営学会

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 6 件)

白井康之; 森田裕之; 中元政一; 高嶋宏之, "共進化的アプローチに基づく"らしさ"の数値化に関する研究 (<特集> データ解析コンペティション: リテールマーケティングの新潮流)", オペレーションズ・リサーチ: 経営の科学, 61, 2, pp.80-87, 2016, 公益社団法人日本オペレーションズ・リサーチ学会(査読有り)

森田裕之, "実践的なデータマイニング教育と学生の獲得能力に関する研究 (<特集> 情報教育の新展開)", 日本情報経営学会誌, 36, 2, pp.51-62, 2015, 日本情報経営学会(査読無し)

中元政一; 高嶋宏之; Cheung, Stephane; 白井康之; 森田裕之, "商品分類と定価推定に関する商品特性の分析方法 (特集 データ解析コンペティション)", 経営システム, 25, 3, pp.164-170, 2015, 日本経営工学会(査読有り)

白井康之; 森田裕之; 中元政一; 小山聡, "人気感度と先行ポイントを利用した顧客セグメント化とその応用 (論文・事例研究, <特集> データ解析コンペティション: インフォメディアリ・データの分析)", オペレーションズ・リサーチ: 経営の科学, 59, 2, pp.88-95, 2014, 公益社団法人日本オペレーションズ・リサーチ学会(査読有り)

Shirai, Yasuyuki; Morita, Hiroyuki; Nakamoto, Masakazu; Oyama, Satoshi, "Analysis Using Popularity Awareness Index, Recency Index and Purchase Diversity in Group Buying", Industrial Conference on Data Mining, pp.100-114, 2014, Springer (査読有り)

MORITA, Hiroyuki; Maheo, Arthur, "Classification Model Using Contrast Patterns and GRASP", Journal of Information Assurance & Security, 9, 5, 2014, pp.235-243 (査読有り)

〔学会発表〕(計 12 件)

Mizuki SAKAI, Hiroyuki MORITA, "Study on a model that identifies moving direction using BLE Devices", 6th International Conference on Social Sciences, Business, Technology and Management, March 18-19, 2017, Seoul, South Korea

Riko KUNIMOTO, Hiroyuki MORITA, "Forecasting Web Page Views Using Contrast Patterns in Fashion Coordinate Data", 6th International Conference on Social Sciences, Business, Technology and Management, March 18-19, 2017, Seoul, South Korea
Cheung, Stephane; Shirai, Yasuyuki; Morita, Hiroyuki; Takashima, Hiroyuki; Nakamoto, Masakazu; Ip, Edward Hak-Sing, "Application of Hidden Markov Model to Analyze Enthusiasts' Dynamics of a Lifestyle Brand", 49th Hawaii International Conference on System Sciences (HICSS), pp.1557-1566, IEEE, January 5-8, 2016, Hawaii, USA

Nishiguchi, Mao; Morita, Hiroyuki, "Predictive Model Considering Coverage of Transaction by Contrast Patterns", 11th International Conference on Knowledge Management, November 4-6, 2015, Osaka, Japan

Cheung, Stephane; Shirai, Yasuyuki; Morita, Hiroyuki; Nakamoto, Masakazu, "Understanding Changes in Customer Purchase Behavior: Study of Attenuation Model for Multiparametric Purchase Preferences", Data Mining Workshop (ICDMW), 2014 IEEE International Conference on Data Mining, pp.274-283, IEEE, December 14-17, 2014, Shenzhen, China

森田裕之, "実践的データマイニング教育と学籍の能力向上に関する研究", 情報経営学会 2014 年秋季研究発表会, 2014.11.7-10, 沖縄県石垣市

西口真央; 森田裕之, "利用パターン数の削減を可能とするパターンを利用したクラス分類予測モデルの提案", 経営情報学会 全国研究発表大会要旨集 2014 年秋季全国研究発表大会, pp.105-108, 一般社団法人 経営情報学会, 2014.10.25-26, 新潟県新潟市

白井康之; 森田裕之; 中元政一, "購買選好度減衰曲線を用いた選択多様性解析とその応用 (特別セッション ビッグデータとマーケティング分析 (2))", 日本オペレーションズ・リサーチ学会秋季研究発表会 アブストラクト集, pp.234-235, 公益社団法人日本オペレ

ーションズ・リサーチ学会, 2014.8.28-29, 北海道札幌市

Shirai, Yasuyuki; Morita, Hiroyuki; Nakamoto, Masakazu; Oyama, Satoshi, "Analysis Using Popularity Awareness Index, Recency Index and Purchase Diversity in Group Buying", Industrial Conference on Data Mining, pp.100-114, Springer, July 16-20, 2014, St.Petersburg, Russia

森田裕之; 白井康之; 中元政一, "エントロピーを応用したフラッシュマーケティングにおける多様性を考慮した分析", 経営情報学会 全国研究発表大会要旨集 2013 年秋季全国研究発表大会, pp.329-332, 一般社団法人 経営情報学会, 2013.10.26-27, 兵庫県神戸市

Morita, Hiroyuki; Nishiguchi, Mao, "Classification Model using Contrast Patterns", ICEIS, pp.334-339, July 4-7, 2013, Angers, France

NAKAYAMA, Yuji; ISHIGAKI, Tomonori; ARAKI, Nagateru; MORITA, Hiroyuki; KONDO, Masashi", "Teaching Business Analytics to Undergraduate Students with R AnalyticFlow", EURO2013, July 1-4, 2013, Rome, Italy

〔図書〕(計 0 件)

〔産業財産権〕

出願状況 (計 0 件)

取得状況 (計 0 件)

6. 研究組織

(1) 研究代表者

森田裕之 (MOIRTA, Hiroyuki)

大阪府立大学・人間社会システム科学研究科・教授

研究者番号: 80295732

(2) 研究分担者

なし

(3) 連携研究者

なし

(4) 研究協力者

なし