

# 実データを活用した問題解決型データマイニング教育

森田 裕之<sup>1,a)</sup> 荒木長照<sup>1,b)</sup> 近藤真司<sup>1,c)</sup> 中山雄司<sup>1,d)</sup> 樋口友紀<sup>1,e)</sup>

**概要:** データマイニングを総合的な問題解決型の教育プログラムとして実施するには、教材としてのデータ準備、必要な学習内容の選択、評価方法、そして実際のマイニングによって得られた結果の実践など様々な課題が存在する。我々は、これらの課題を解決すべく、分析可能な実際の POS データを準備するとともに、提案内容を実践する取組を 10 年程度の時間をかけて教育プログラムとして整備してきた。そしてこの実施内容については、プログラム独自でアンケートを調査を受講学生に対して実施し、評価を行なっているところである。本稿では、本教育プログラムについて説明すると共に、これまで得られたアンケート調査の結果などから、プログラムについて評価を行い、今後の課題について明らかにする。

**キーワード:** POS データ, 問題解決型教育, データマイニング

## A problem based learning program about data mining to use a real POS data

MORITA HIROYUKI<sup>1,a)</sup> ARAKI NAGATERU<sup>1,b)</sup> KONDO SHINJI<sup>1,c)</sup> NAKAYAMA YUJI<sup>1,d)</sup>  
HIGUCHI YUKI<sup>1,e)</sup>

**Abstract:** It is difficult to implement an educational program which is problem based learning about data mining. Because it is not so easy to prepare a real business data, to select necessary educational contents, and to implement a business solution which students propose in real. We have arranged the environments for the data mining and have implemented the educational program about that for the last ten years. And we obtain information by means of questionnaires from students who experience the program and we evaluate our program from an analysis of the data in this paper.

**Keywords:** POS data, Problem based Learning, Data mining

### 1. はじめに

ICT(Information and Communication Technology)の発達により、「ビッグデータ」の関心が高まっている。そして、これを活用するデータマイニングを企業経営に取り入れようとする動きが活発になってきている。データマイニングという言葉が、使われ始めてからは、すでに 20 年ほ

どの時間が経過しており、その間、企業での利用が注目されていたが、順調とはいえない状況にある。これにはさまざまな理由があるが、その 1 つは、せっかく開発されたアルゴリズムが十分に活用されていないことが 1 つの要因である。そしてその活用が困難である理由は、企業に必要なデータマイニングが問題発見からソリューションを提案して実践するまでの一連の PDS サイクルであり、これを実践可能な人材が少ないことによる。また、これを十分に教育できる体制を大学としても提供することが難しく、十分な人材を輩出できていないこともその要因として存在している。我々は、経営学やマーケティングをベースとしながら、必要なマイニング技術を活用できるように、データマイニングを問題解決型のトータルプロセスとして捉え、これ

<sup>1</sup> 大阪府立大学  
Osaka Prefecture University, 1-1 Gakuenmachi, Sakai, Osaka  
599-8531, Japan

a) morita@eco.osakafu-u.ac.jp

b) araki@eco.osakafu-u.ac.jp

c) kondo@eco.osakafu-u.ac.jp

d) nakayama@eco.osakafu-u.ac.jp

e) higuchi@eco.osakafu-u.ac.jp



図1 本プログラムでのデータマイニングの考え方

を習得できような教育プログラムとなるように、複数の教員で協力し、10年ほどの時間をかけて整備してきた。具体的には、まず、大学で営業している生活協同組合の協力を得て、購買や食堂のPOSデータを提供いただき、これを学生に提供している。そしてこのデータを学生に提供して分析するとともに、その提案を一部実践している。このデータを利用して、分析を行うことは、実際のPOSデータを利用できるという利点だけでなく、POSデータを構成するトランザクションイベントが観察できる範囲にあるという利点もある。これによって、データだけでは判明しない問題点も実店舗を観察することでカバーすることができると共に、必要であれば、利用者へのアンケートなども実施することが可能である。この環境のもとに、授業において共通の課題テーマを毎年設定し、学生に提示する。学生は、その課題テーマに対する問題を発掘し、データを分析し、様々な観察や分析を行いながらチームで最終的な、ビジネスソリューションを提案する。提案されたソリューションを、授業を実施しているゼミ内で相対的に評価すると共に、複数の教員のゼミ間で集まって最終的な優劣を競い合う。その後、優秀な提案の上位2チームについては、提案内容を実際の店舗で実施して、その提案内容の妥当性を客観的に評価できるプロセスまで行なっている。もちろんこれらのプロセスを実行するには、学生に提示する課題テーマの選択、ソリューションとしてまとめ上げるまでの教育及び指導方法、そして提案を実施する方法など、いろいろ検討しなければならない問題が存在している。また、学生の達成感や能力獲得度合いの格差などの問題も存在している。

以下では、最初に我々の取組に至った背景、取組の概要を説明し、その後、独自に取得している本教育プログラムに関するアンケート調査の結果から、プログラムの評価と今後の課題について議論する。

### 生協データ解析コンペティションの流れ

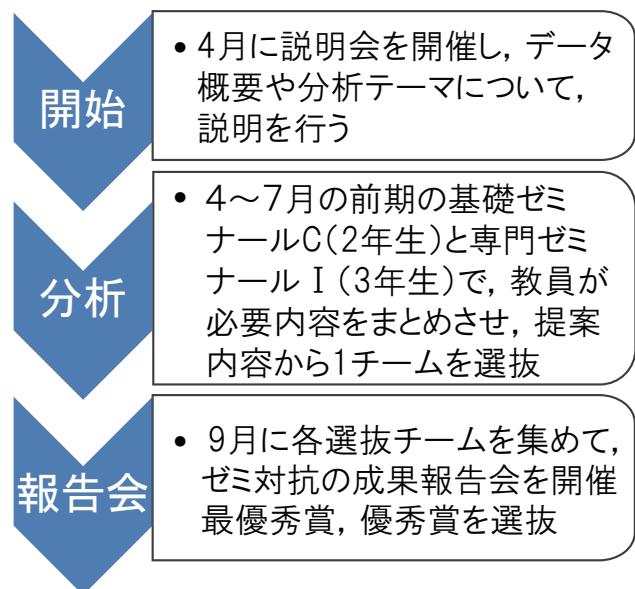


図2 生協コンペの流れ

## 2. 本教育プログラムの概要

本プログラムでは、データマイニングを全体的な問題解決プロセスと考え、問題発見から実践までの広義な意味で考えている。したがって図1のようなイメージとなる。これは前述のように、実際にビジネス現場においてデータマイニングを実践できる人材を教育したいからである。

これを実践するために、本プログラムでは2つの大きなイベントを用意している。1つは「生協データ解析コンペティション(以下、生協コンペと呼ぶ)」で、もう1つは「生協チャレンジ」である。前者は、テーマにしたがってデータを利用しながら、ビジネスソリューションとして、実際

的な提案をするまでを行うイベントであり、後者は、その中の優秀な提案を実際に店舗で実践するイベントである。

生協コンペでは、対象学生を主として2年生（一部、3年生\*1）基礎ゼミナールCという2年生前期に提供される授業で実施している。まず4月ごろ全体の説明会を実施して、共通のデータを提供した上で、データの説明と分析テーマを説明する。分析テーマは、毎年異なるものを設定するようにしている。最近では、実践する際の商品廃棄ロスの問題などを考慮して、購買部門のお菓子の販売についてのテーマを設定している。これは、毎年事前に生協の販売担当者と関係教員がミーティングを行なって、彼らが直面している問題のうちテーマとして可能であるものを教員側で判断して設定するようにしている。授業では、テーマと連動して、必要な内容を学生に教育しながら、分析結果をまとめることができるようにガイドしていくことになる。主たる対象が2年生ということもあって、1年次の基礎的な講義は受けているものの、もちろん分析結果をまとめるのに十分な能力を有しているとは言い難い。したがって、多くの内容を教育したいと思う一方で、半期の15回の授業内で結果を出せるようにしなければならないという時間的な制約も大きい。また当初は、各教員の判断で内容や教育方法を検討していたが、回数を重ねるにつれ、店舗の視察を兼ねた観察力を養う方法、分析の基礎的な内容、そして何らかの分析手法については、少なくとも簡単な教育を行う必要があり、それに加えて、何らかのチームで作業する訓練も必要であると認識されている。

一方、学生は授業を受けながら、テーマに沿った問題の発掘を、店舗の観察とデータの分析から行う。例えば、店舗での売上最大化というテーマを設定したとしても、どのようにして最大化するかは必ずしも1つではない。単一の分類に着目する方法や、複数の分類の併売を促進する方法など、最大化する方法も様々である。

次に問題を解決するのに必要なモデルを作成する。そのとりかかりとして、何が目的変数で、何が説明変数となるのかを明確にする必要がある。通常、これは教科書を使った授業では、予め設定されていて、これを使って必要なモデルを作成することが多い。しかし実際には、何を目的変数とすべきで、それを説明可能と思われる変数が存在しているか、または存在しているデータから作成することが可能であるのか、といったことを考慮することは、大変重要なプロセスとなる。これはおそらく、モデルがどのように作成されるのか、つまり手法をある程度知っていなければイメージのわきにくい作業なのかも知れないが、現実問題を解決するには、専門知識を持つ担当者とともに考えるべき重要なプロセスである。

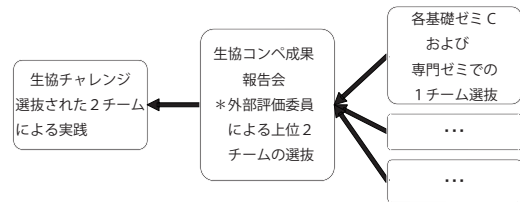


図3 2つのイベントと選抜の流れ

これらの変数を識別した後で、実際に分析のモデルを作成する。もちろん変数識別および選択と、モデルの予測結果は連動するものであるため、必要に応じてフィードバックプロセスを必要としている。モデルは、本来なら複数の異なるモデルを習得したうへでもっとも適切なモデルを選択できるようにできるような環境を整えたいところであるが、この部分については、時間制約と、全体的なソリューションを提案できることを優先しているため、現実には1つのモデルを教えることが限界である。ただし、どのモデルを学生に教えるかについては、各教員の専門としているものを教育される場合や、テーマに従って毎年変更される場合などがあって特に統一しているわけではない。

最後に計算された結果を正しく解釈して、そこからビジネスソリューションに変換するプロセスが必要である。この際、あまり教員がガイドせずに学生にアイデアを考えさせると、値引きやポイント付与などの安易なプロモーションに陥りやすい。また、我々は生協チャレンジも想定しているので、できるだけ実現可能かつ客観的に解釈可能なプランを提案できるような指導を心掛けている。

以上のようなプロセスを経て、基礎分析、問題発掘、モデル作成など、それぞれのタイミングでチームによる内容のプレゼンテーションを実施しながら、最終的な提案としてまとめ、ゼミ内で選抜を行う。ゼミ内の選抜においては、指導教員が単独で評価して決定する場合と参加者が全員で、自チーム以外のチームを評価してその集約した結果によって選抜チームを決めるケースとがある。これによって選抜されたチームは、最終の成果報告会で、各ゼミから選抜されたチーム間でその提案内容を競い合うことになる。この成果報告会では、指導している教員は敢えて審査には加わず、外部から企業家、専門家、そして学内のチームを出していない教員が審査委員となって、選抜チームから上位の2チームを決定する。中には、現実的な側面からかなり厳しいコメントを頂戴することもあるが、それも学生にとっては良い刺激となっている。

生協チャレンジでは、店舗内における対象商品の受発注と商品陳列、そして広告などを学生が担当する。またそれ以外に付加的なアクション（たとえば福引など）を実施する場合は、それも担当する。レジについては、他の商品も存在しているので生協職員に担当してもらうことにしている。商品受発注については、学生はほとんど経験がなく、

\*1 3年生は専門ゼミナールという卒業研究の前段階のゼミナールに属して参加する

## 生協チャレンジの流れ

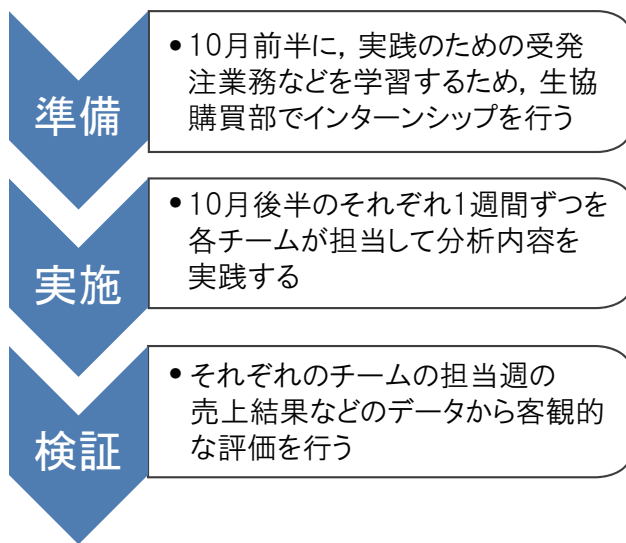


図 4 生協チャレンジの流れ

また万一間違えて発注すると大変なことになる。さらに、当該店舗では在庫をほとんど持つことができないため、それを考慮した受発注と商品陳列が必要となる。これらのことを自分達だけで実施できるようにするため、10月の前半2週間は、当該店舗でインターンシップを実施している。その後、後半の2週間は1週間ずつ担当して各チームが提案を実践している。生協チャレンジの評価については、当初設定したテーマを客観的に評価するようにしている。例えば、売上高を最大化するという目的であれば、売上額が多いほうが優勝チームとしている。これについては、毎年11月に学園祭があり、その中の1つのイベントとして、生協チャレンジ報告会を企画し、そこで実践内容について学生が報告した後、結果発表と表彰式を実施している。

### 3. 授業後のアンケート

大学全体として授業アンケートが存在しているが、本プログラムについては、独自にプログラムの評価を行なうためのアンケートを別途実施している。アンケート自体は、まだ開発途上であり、今後も調査項目の改善の必要があるが、現状取得できている2008年度と2010年度のアンケートについて以下で分析を行なう。それぞれのサンプル数は77件と15件である。調査項目のうち表1の25項目については、両方の調査で共通した項目で2010年度の調査では、これに能力アップに関する主観的な評価とこのようなデータマイニングの取組への興味度の変化を追加して調査している。

質問項目は、本プログラムでの習得を目標としたチーム力、観察力、分析力、そして交渉力が含まれるような項目を設計しており、5件法で5がもっとも”強くそう思う”ように質問している。

表 1 本教育プログラムで実施しているアンケートの項目

No	質問内容
Q1	この授業のようにグループ単位で作業することは楽しい
Q2	この授業を受けて自分に足りない知識が何かわかった
Q3	もっと PC (SPSS など) について勉強したくなった
Q4	最初わからなかったことがある程度解決できた
Q5	生協以外のコンビニなどでも商品の売れ方が気になる
Q6	データを加工して問題を発見できた
Q7	生協の販売現場にいて問題を発見できた
Q8	教員による講義内容をヒントに問題発見ができた
Q9	グループで話し合っ問題発見ができた
Q10	自分で勉強したことから問題発見ができた
Q11	他の授業より担当教員に対して質問をしたと思う
Q12	グループ (G) のメンバーとよく話し合ったと思う
Q13	学習を進めるために生協の方に質問しに行った
Q14	G メンバーに自分の得意なことを教えたことがある
Q15	G メンバーから自分の不得意なことを教えてもらった
Q16	新しい PC ソフトの使い方を身につけた
Q17	この授業のことで Web を使って調べた
Q18	この授業のことで本 (参考書など) を読んで調べて
Q19	自分を含む消費者の意思決定に関することを学んだ
Q20	企業の意味決定に関することを学んだ
Q21	授業が始まったころに比べて自信をもって発表できる
Q22	発表スライドの作成法がわかった
Q23	発表の際の注意事項を3つ以上いえる
Q24	発表を聞く際に適切に評価メモをとれる
Q25	発表者に対して適切な質問ができる

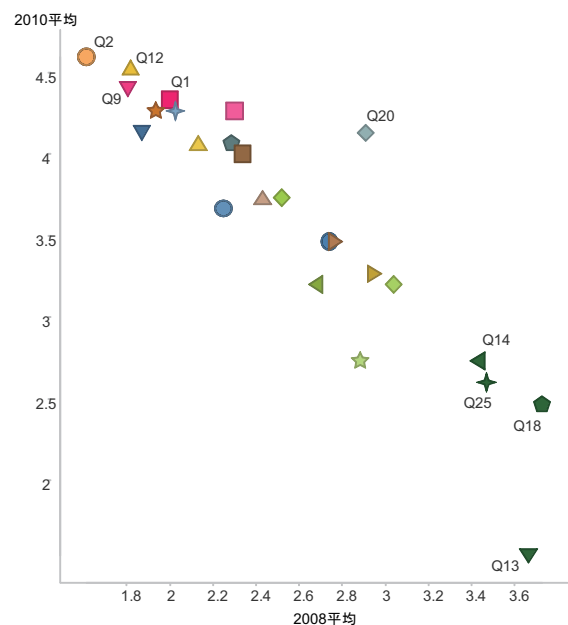


図 5 2年間の各質問項目の平均値比較

図5は、25の質問項目の各項目の平均値を、各年の次元でプロットした図である。図の見方としては、右上にプロットが存在すれば、両年とも評価がよく左下はその逆となる。また右下は2008年は評価が良く2010年は評価が低く、左上はその逆であると考えられる。まず左下に存在する項



目がないことから、この2年間において、両方とも極端に悪かった項目はないことがわかる。しかし一方で、右上の項目も、あまり存在せず、若干、"Q20:企業的意思決定に関することを学んだ"という項目が該当していることがわかる。これは、今まで実際の問題解決的な課題を考える授業を受けていなかったこととの違いによってそのように感じる学生が多かったのではないかと予想される。次に、比較的右下に存在する点は、次の4つの項目である。"Q13:学習を進めるために生協の方に質問しに行った"、"Q14:Gメンバーに自分の得意なことを教えたことがある"、"Q18:この授業のことで本(参考書など)を読んだりして調べた"、"Q25:発表者に対して適切な質問ができる"。また、比較的左上の点は、"Q1:この授業のようにグループ単位で作業することは楽しい"、"Q2:この授業を受けて自分に足りない知識が何かわかった"、"Q9:グループで話し合っって問題発見ができた"、"Q12:グループ(G)のメンバーとよく話し合ったと思う"。である。これらを解釈することは、なかなか難しいが、1つの原因は、学年よる学生の違い、指導する教員の違い、テーマの違いなど環境が大きく異なっている状況において、2008年は比較的半数に近い調査で、2010年は、自主的に回答した学生による回答で、かなり傾向がことなる回答になっていることが考えられる。その前提から、結果を解釈すると2008年の回答は、比較的自主性と能動的な学習態度が見られるのに対して、2010年の回答は、チームとして活動した学生の回答が強調されているように思われる。

次に因子分析を行なって、どのような潜在因子が確認されるかをみてみることにする。図6は、7つの因子が抽出された状態で、回転後の累積寄与率は57.9%である。因子としては、前述のチーム力、観察力、分析力、そして交渉力をより細分したものが出現している。"授業からの学習達成感"は、チーム力をベースとした複合因子で"データ分析からの学習達成感"は分析力を、"コミュニケーション"と"教師とのやり取り"は交渉力を"現場問題発見"は観察力を、そして残りの2つは能動的な学習を表していると考えられる。

これらのうち年度ごとに各因子得点を平均して比較し、有意な差が確認されたものが、図7である。図より、2008年と2010年を比較すると、コミュニケーションと自主学習は得点が悪化しており、2つの学習達成感に関しては改善していることが確認できる。しかし、これは我々の努力によって値が改善したり、また、何か努力不足によって得点が悪化したということではなく、前述の環境と学生の層の違いによるものではないかと思われる。

このことを確認するために、2つの学習達成感の得点を軸として2次元に各サンプルプロットしたのが、図8である。図より、単に平均が異なるだけでなく、サンプルが明確に2つのグループに分かれていることがわかる。

質問番号	因子						
	授業からの学習達成感	データ分析からの学習達成感	コミュニケーション	プレゼン参加意識の向上	自主学習	教師とのやり取り	現場問題発見
Q12	.777	.422	.080	-.127	-.086	.057	-.010
Q1	.725	.209	.048	-.117	-.174	.275	-.084
Q2	.717	.358	-.069	.062	-.160	-.010	-.126
Q5	.658	-.025	-.080	.209	.162	.219	.096
Q9	.633	.607	-.055	.054	-.104	.119	.070
Q22	.595	.280	-.113	.187	.011	-.007	-.042
Q3	.583	.327	-.029	.132	-.054	.173	-.304
Q19	.572	.172	-.232	.399	.045	-.094	.245
Q21	.523	.171	.171	.350	.019	.387	-.002
Q8	.506	.288	-.115	.100	-.052	.410	.188
Q20	.477	.197	.197	.153	.084	-.028	.191
Q16	.385	.681	-.118	.038	-.101	.125	-.044
Q6	.384	.679	.084	-.029	.070	.034	.279
Q10	.112	.599	.009	.222	.141	.108	.105
Q4	.427	.594	-.085	.099	.208	.240	-.086
Q14	-.057	.054	.816	.167	.103	.010	-.002
Q15	.468	-.109	.528	.133	-.096	-.188	.070
Q13	-.229	-.190	.404	-.055	.221	-.056	.386
Q24	.070	.229	.111	.656	-.058	.137	.117
Q23	.173	.029	.064	.422	-.059	-.034	-.024
Q25	-.170	-.125	.310	.412	.161	.320	.049
Q18	-.042	-.004	.091	.029	.809	-.009	.078
Q17	-.014	.253	.043	-.294	.459	.144	.079
Q11	.261	.294	-.108	.042	.079	.607	.085
Q7	.051	.183	.030	.101	.081	.119	.733

主因子法+バリマックス回転による

図6 因子分析で抽出された因子

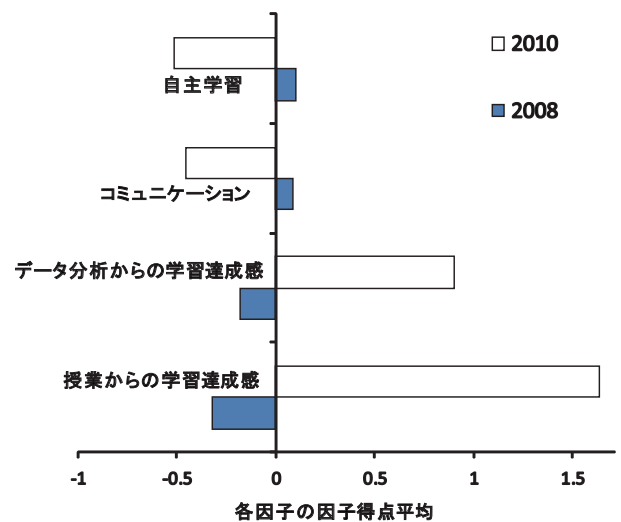


図7 因子得点の平均比較

次に2010年度の調査のみが対象となるが、能力アップに関する主観的な評価とデータマイニングの取組への興味度の変化を追加で調査しているため、それに関して、追加的に質問項目との関係について分析する。

まず能力アップに関する主観的な評価を、授業受講前の分析能力を100%として、どのくらいアップしたと感じているかを質問したのが、図9の上図である。回答は4段階で、"変わらない"、"20%ほどアップした"、"40%ほどアップした"、"50%ほどアップした"と感じるかを質問している。結果から回答した大多数の学生は、能力アップを感じており、40%以上の能力向上を感じている学生が半分以上

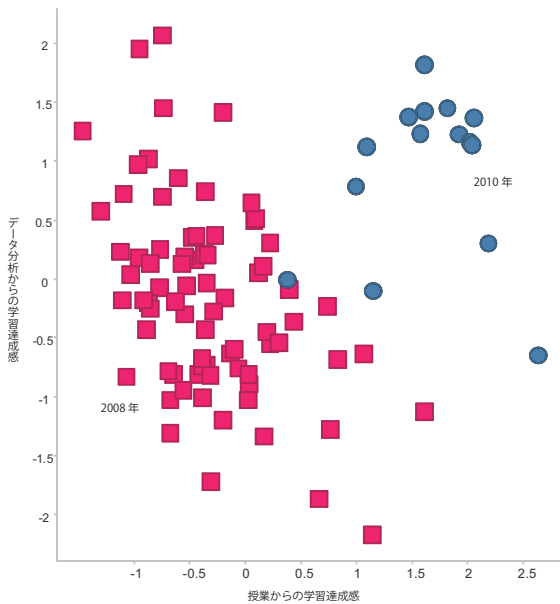


図 8 第一と第二因子の次元でのプロット

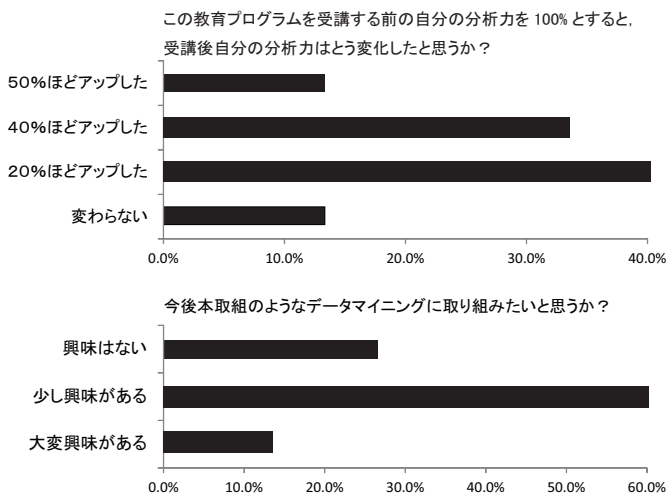


図 9 学生からみた能力の向上と関心の度合い ([1] より引用)

いることがわかる。もちろんこれは、実際にそうなっているかどうかの客観的な判断は難しいが、それだけ学生が努力したと感じ、それによって能力も向上したと感じていることの証左ではないかと考えられる。また、プログラムのようなデータマイニングの取組に関する興味の程度を3段階で質問したのが、図9の下図である。程度としては、少しというのが最も多いが、このプログラム自体が、データマイニング教育のステップアップするための第一段階と位置付けているため、興味を持ち始めるかどうかについては、大きな違いであると考えており、満足いく結果であると思われる。

次に、これらの能力アップと関心の度合いは、Q1~Q25のどの質問項目の回答と関係しているのかを決定木分析を用いて調べてみることにする\*2。まず、図10の、能力向上

\*2 決定木分析は、数理システムの VMStudio を用い分割方法は InfoGain Ratio, 各最大分岐数は 2 と設定して、以下の分析を行

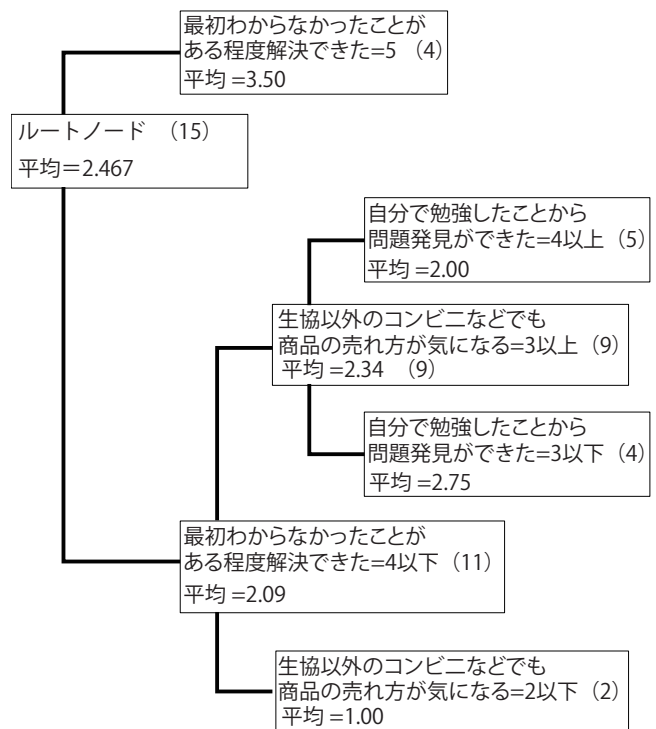


図 10 主観的な能力向上とそれをもたらす要因

を目的変数とした分析では、目的変数を数値属性として与えてモデルを作成した。図中の四角のノードは分岐点または葉ノードを表しており、ノードの内の記述は、分岐条件、括弧内の数値が該当サンプル数、そして平均は、1~4の回答数値の平均を表している。したがって平均=4であれば、そのノードはすべて、"50%ほどアップした"と回答したことを意味している。決定木から結果を考察すると、何か疑問に思ったことを解決できるという成功体験を経験させることが、能力アップを感じていることに一番大きな影響力を持っていることがわかる。これは当たり前のことであるようにも思えるが、我々としては、つい難易度の高いハードルをかかげてしまいがちであり、それは能力アップを感じる機会を損ねてしまうことにつながるという警鐘であるとも理解できる。すべてのハードルを下げることは難しいが、容易なものや難しいものを用意して、学生のレベルに合わせて使い分けることが重要ということであろう。また、第2分岐と第3分岐から、プログラム内容に興味を持たせることは重要であるが、自発的な学習がその成果に結びつくかどうかはあまり関連がないと解釈できる。これは目的関数が主観的判断に基づいている影響であると予測でき、逆に目的関数を実際の各サンプルの成績などを利用すれば結果は異なるかもしれない。

最後に、目的変数を興味度合いに変更して、同様の決定木分析を行なう。興味度は3段階で、1:興味ない、2:少し興味がある、3:強く興味がある、という回答を、3つのクラス分類問題として目的変数を設定し、計算を行った。各

なっている

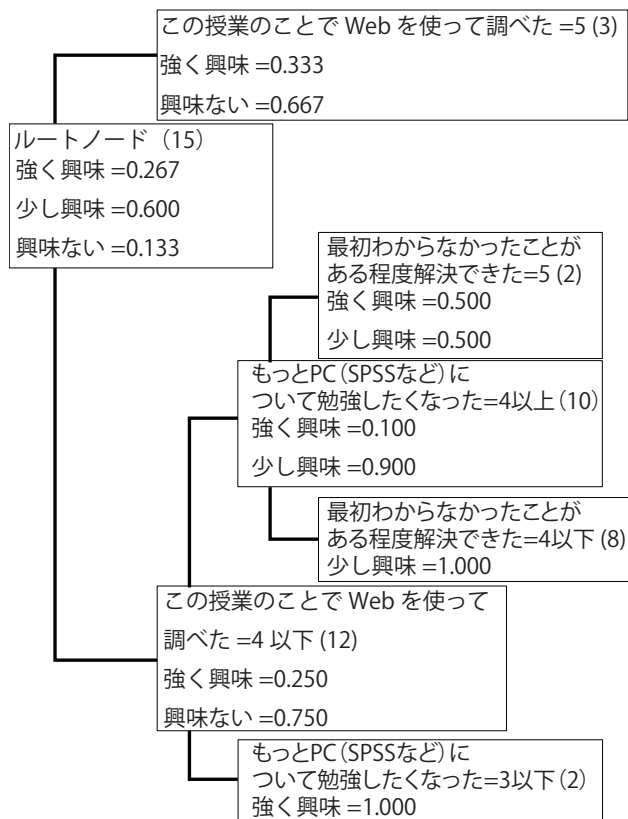


図 11 学生のプログラムに対する関心に影響する要因

ノード内の分岐条件と該当人数は、図 10 と同様であるが目的変数は、各ノードの 0 より大の各クラスの構成比率を示している。したがって、例えば図 11 の一番上のノードは、「この授業のことで Web を使って調べた」を「大変そう思う」と回答した 3 サンプルのうち、33.3%は「強く興味」をもっているということを示している。図 10 がその結果を表しているが、解釈はなかなか難しい。第一分岐では、自分で Web を使って調べたとしても興味を持った学生となかった学生が混在している。しかし、第 2 分岐を見ると、ソフトウェアの学習意向が強いほうが、関心を持っている学生数が多いことから、その因果関係は明らかではないが、何か具体的な課題で知的好奇心を刺激することが重要なのではないかと考えられる。

#### 4. おわりに

本稿では、実データを活用したデータマイニング教育プログラムの 1 つの実践例を紹介すると共に、プログラムに関連した学生へのアンケート調査を踏まえて、状況を確認した。まだまだ、細かな教育方法については考慮しなければならない点は存在しているが、問題解決型の実践的なデータマイニングを教育する教育プログラムとしての内容とその方法は 1 つの完成形に近い形式になっているのではないかと考えている。アンケート結果の分析と我々が感じている問題点を最後に議論する。アンケート分析からわ

かった点をまとめると、以下ようになる。

- (1) 環境や学生の違いなどによっても学習による習熟度合いや観点が異なっている。
- (2) チーム力、交渉力、観察力、分析力といった獲得能力の観点は因子分析からも確認されている。
- (3) 因子分析で識別された学習達成感は、学生によってもかなり異なる可能性がある。
- (4) 主観的ではあるが能力アップや興味を持つことについては、本プログラムによって達成されているようだ。
- (5) 能力アップ感や興味を持つきっかけとしては、学生が達成可能なレベルを考慮した課題設定が必要であるとともに、そのバリエーションも考慮する必要がある。

これらの改善に関して、学生の差異などは指導教員が感覚で判断せざるを得ない部分もあるが、関係教員で議論した結果、大きく 2 点については検討・試行すべき点が明確になった。

1 つは、これまで我々はデータマイニングの教育プログラムであることを念頭に置いていたため、無意識に分析手法や分析方法に焦点を当てすぎていた傾向があったかもしれないという点である。これは、限定された講義期間内で一定の結果をまとめなければならないため、やむを得ない点もあるが、それによってあまり分析が得意でない学生に対しては、達成感が十分に得られない結果をもたらした可能性がある。この点の改善については、もう少し観察力を重視するようなプロセスを採用する改善が考えられる。具体的には、販売現場がすぐそこにあるわけだから、これまで学生の自主性に任せていた店舗の視察を授業中にも採用し、データ以外からわかること、そして観察しなければならぬことを明確にするプロセスを追加することである。もちろん、これも単に視察に行くだけでは、あまり意味をなさないため、達成感を得られるようなチェックポイントを、視察前に具体化して学生が何かを成し遂げた実績を残せるような工夫を教員側でも行なう必要がある。これによって、これまで結果評価に偏りがちであった評価方法も、少しプロセスを評価する要素を追加できるのではないかと考えている。

もう 1 つの点は、全体として上記のようなチェックポイントを細かく設定して、成し遂げる課題についての部分的な達成感を重視する方法の重点化である。これまで、教員によって差があるものの、このような方法は採用されてきたが、アンケート結果からも、自分にできないこと・わからないことが解決されることによる達成感や充実感は予想以上であった。一方で分析のプロセスは、基礎分析やモデル設計、マイニングそして解釈といったプロセスを一通り実行するればよいわけではなく、どこかのプロセスまでいっても、不十分であれば必要な部分まで戻るといった繰り返し作業が必要となる。また、それが必要であることを教えることも、重要な点であると考えられる。そのため、各ブ

ロセスにおける有効なチェック方法を型どおりに設定して効果を生むのであれば良いが、その方法については、今後も試行錯誤を繰り返して検討する必要があると考えている。

最後に、アンケート調査について考察を行なう。調査項目については、ある程度十分であると考えられるが、”Q23：発表の際の注意事項を3つ以上いえる”については、質問の重複感、およびこれまでの分析からもそれほど重要であるとは言えないので、削除してもいいかもしれない。また、学生の違いが今回の分析では見られたが、これが環境に依存するものなのか、学年の違いなのか、または成績の違いによるものなのか興味深い点である。可能であれば、客観的な能力に関する変数を導入することでより明確にできないかを検討したい。

また、アンケート調査についてはまだ開始してから数年であり、そして1年で1回分しか回収できないため、データ量が十分であるとは言えない。これも長期間にわたって同じ調査を実施することで、今後得られる様々な教育上の示唆があると考えられる。アンケートの設計自体は上記のような微修正が必要であるが、大枠については十分であると考えており、今後、データを蓄積することでそこから得られる結果を取り入れて、より良い教育プログラムとなるように改善を重ねていきたいと考えている。

謝辞 大阪府立大学生協同組合の皆様には、データ提供ならびに本プログラムの実施において多大な協力を頂いている。本プログラムは、平成20～22年度の文部科学省質の高い大学教育推進プログラムで、「販売現場に密着した問題発掘型スタディーズ」として採択されたものである。

#### 参考文献

- [1] 森田 裕之, 荒木 長照, 近藤 真司, 中山 雄司, 樋口 友紀, 『大学生協 POS データを活用したデータマイニング教育』, PCカンファレンス2011 予稿集, (2011)
- [2] Hiroyuki Morita, Nagateru Araki, Shinji Kondo, Tomonori Ishigaki, Yuji Nakayama, "AN ENHANCED PRACTICAL PROGRAMMEME ON DATAMINING EDUCATION", Proceedings of the 2nd International Conference on Computer Supported Education, 2010, Valencia, SPAIN