

## コンビニにおける新商品発売時の売上データマイニング - 目的変数と説明変数の導出 -

菅野 憲明\*, 佐々木雄大†, 高山 毅\*†, 池田哲夫\*†

コンビニエンス・ストア（通称「コンビニ」）における新商品の売れ行き予測に関する学術的な検討が遅れている。著者らはこれまでに、市販の POS データを用いて「おにぎり」と「カップラーメン」の新商品の売れ行き予測に関する予備的な検討として、特定店舗で成立する相関ルールの抽出を行なった。本稿では、店舗を問わず成立する一般的知見の導出を目指す。具体的には、コンビニ会社 A 社へのヒアリングを通じて、新商品投入後二週目の売れ行き予測に対するニーズがあることを述べる。そして、これを目的変数とした場合の説明変数として有効な変数を、A 社の協力の下、データマイニングにより導出する。これらの説明変数と目的変数を用いて、店舗ごとにルール抽出を行なうと有効なルールが得られることが確認された。

### Sales Data Mining When New Items are on Sale in Convenience Store -Derivation of a Target Variable and its Explanatory Variables-

Noriaki Kanno\*, Takehiro Sasaki†, Tsuyoshi Takayama\*†, and Tetsuo Ikeda\*†

Academic analysis of sales forecast for new items on sale in convenience store is insufficient. As a preliminary analysis, the authors have derived some association rules for every shop concerning demand of new "onigiri" and "cup ramen" by utilizing POS data on sale. In this paper, we try to find some general knowledge independent of a shop. Concretely, based on a hearing from a convenience store company A, we describe there exists a need for sales forecast in the second week for a new item. We adopt it as a target variable, and derive its effective explanatory variables by the use of data mining technique, based on the cooperation of the company A. We have confirmed that we can obtain some effective association rules, if we adopt these target variable and explanatory variables and perform distinct data mining for each shop.

#### 1. はじめに

小売業において売れ行き予測は重要である。

---

\* 岩手県立大学 大学院 ソフトウェア情報学研究科  
Graduate School of Software and Information Science,  
Iwate Prefectural University

† 岩手県立大学 ソフトウェア情報学部  
Faculty of Software and Information Science, Iwate  
Prefectural University

そのため、小売業においては、新商品の発売時には事前のマーケティング調査などにより売れ行き予測が行われている。しかし、事前調査の情報のみからの確に発売後の結果を予測するのは容易とはいえない。

売れ行きのデータが POS 等によって電子化されている場合、売れ行き予測にデータマイニングを用いることが有力視される。コンビニエン

ス・ストア(以下「コンビニ」と略)は一般に売り場面積がスーパーマーケット(以下「スーパー」と略)よりも小さめであることが少なくなく、売れ行きが不調な商品を置く余裕がないことから、より精度の高い売れ行き予測が求められると予想される。

コンビニでの新商品の売れ行き予測に関する学術的な分析は、現在までのところ充分になされていない[1]。データマイニング分野全体の傾向として、業務先行で企業単位で閉じて知見が蓄積されている可能性もあるが、学術的に整理して後世へ伝えていく意義は、小さいとは言えない[2]。そこで著者らは、コンビニにおける新商品の売れ行き予測の研究を行なっている[3]。文献[3]では手始めに、有料で販売されているPOS データ[4]を購入、分析して、新商品投入後の日数と売れ行きに関する特定店舗での相関ルールの抽出を行なった。ルール抽出には成功したが、購入コストの制約から、

**(問題点1):** データ量が充分とは言えず、抽出したルールの一般性について課題が残されている[5]。

また、文献[8]で指摘されているように、

**(問題点2):** 個々の店舗を対象として、数値を含めて予測式を作ろうとすると、店舗の数と同数の予測式が必要になる。

本稿では、コンビニ会社の協力を得た上で、文献[3]の約 15 倍のデータを準備する。そして、売上の数値をあてる予測式の導出を直接行なうのではなく、文献[6]と同様に、この対象領域において、何を目的変数、何を説明変数とするのが適切なのかに関する分析を行なう。

本稿はこれ以降、次のように構成されている。まず 2 節ではコンビニにおける新商品の売れ行き予測において、コンビニ独自の手法の必要性について述べるとともに、関連研究をまとめる。3 節では目的変数と説明変数のニーズ分析を、コンビニ会社 A 社へのヒアリングを通じて行なう。

そして 4 節では、業務上有用な目的変数をひとつ採用した上で、データマイニングにより、説明変数を導出する。5 節では、ここまでで得られた説明変数と目的変数から店舗単位でルール抽出を行ない、評価する。最後に 6 節で結論と今後の展望を述べる。

## 2. コンビニの新商品の売れ行き予測

### 2.1 コンビニ独自の手法の必要性

スーパーでの新商品の売れ行き予測に関する先行研究は、文献[7]記載の通り種々存在する。しかしながら、コンビニでの新商品の売れ行き予測は、スーパーと違ったアプローチが必要である。その理由として、スーパーとコンビニには以下に示す三点の違いがある。

#### (1) 売り場面積

スーパーと比較してコンビニは一般に売り場面積が狭めのケースが少なくないので、売れないと判断された商品は早期に別の商品に置き換えられてしまう。また、売り場面積が狭めということは、スーパーの場合よりもさらに、各商品の売れ行きを正確に予測して、適切な商品を適量仕入れる必要性が大きくなるといえる。

#### (2) ポイントカード

顧客を特定する方法にポイントカードがある。特定顧客の購入履歴は、売れ行きを予測する上で重要な情報である。スーパーではポイントカードの利用で顧客の特定が可能であるが、コンビニではポイントカードの利用がスーパーほど一般的ではないので、顧客の特定が出来るとは限らない。

#### (3) 特価販売

スーパーでの売れ行きは、商品そのものの人気のほかに、特価販売の有無に多大な影響を受

ける。しかし、コンビニでは定価販売が主流であって、スーパーほど特価販売が一般的ではない。別の見方をすれば、スーパーの場合のように、特価販売によって売れ行きの意図的な制御を行なうことが一般的にはなっていない。

## 2.2 関連する研究

売れ行き予測の現状に関して、商業的な視点から、および学術的な視点からまとめる。

商業的には以下のような事例が見られる。

文献[8]によれば、明治乳業とヴァル研究所の共同開発製品「客ごころ」では2004年6月現在、気象データやイベント情報と過去3ヶ月分の来店数/売上高データを基に、重回帰分析を用いて来店数/売上高の予測をするメニューがある。しかし、何がいくつ売れるかまでは予測できない。

文献[8]によれば、コンビニ大手のファミリーマートは、ISDNを用いて全国4000の店舗に売れ筋分析データや天気予報等のデータを提供している。しかし、売れ行き予測については不完全であり、最終的な仕入れ量の判断は各店長に委ねている。

アイ・ビー・エム社の「INFOREM」は、可変応答法と呼ばれる統計手法を用いて売れ行き予測を行い、在庫数量を考慮して最適な発注を自動的に行なうシステムである[9]。しかしこれは、旧来からの商品の売れ行き予測を行なうものであり、新商品の売れ行き予測には利用できない。

味の素株式会社では、トラッキング(追跡)モデルを用いて新商品の売れ行き予測を行っている[8]。しかしこれは、メーカー側から見た全国的な売れ行き予測であり、コンビニの店舗レベルでの商品単位の売れ行き予測を提供するものではない。

現状では商品の売れ行き予測機能を有する商用システムは、「コンビニの店舗レベルでの商品単位の新商品の売れ行き予測」という需要にこたえる物ではない。

学術的には、国立情報学研究所の論文検索システム NACISIS-IR[13]や情報処理学会、電子情報通信学会、人工知能学会等の論文検索システム(それぞれ[14], [15], [16])で検索しても、当該問題に関する研究は皆無であり、学術的な検討が遅れている。

新製品の売れ行き予測に関して経済学では、「トライアルリピートモデル」というモデルが代表的である。ここでいう「トライアル」とは新規購入者、「リピート」は反復購入者のことである[10]。どの程度の割合で新製品が試行的に購入されるか、その購入者のうちどの程度の割合が再度その新商品を購入するか、という分析から新商品の売れ行きを予測する。しかしながら、このモデルでは売れ行きの予測のために、6週間程度の期間の売れ行きデータが必要になる[11]。

データマイニングを用いてこれより短い期間で売れ行きの予測ができれば、有意義といえる。

## 3. 目的変数と説明変数のニーズ分析

### 3.1 大学からコンビニ会社への打診

まず、今回協力を得たコンビニ会社 A 社提供の POS データを事前分析した結果を踏まえ、以下の三項目を A 社に打診した。

**(打診 1):** 1日単位での売り上げではバラつきが少なくないので、より広い時間幅での累計売上個数を採用して分析を進めること。

**(打診 2):** 平日と土日で客層が異なると推測して、週を「平日の累計売上個数」、「土日の累計売上個数」という変数を設けること。

**(打診 3):** (打診 1~2)の状況を踏まえつつ、  
目的変数: 「第4週の累計売上個数」  
説明変数: 「第1週の平日の累計売上個数」、「第1週の土日の累計売上個数」、「第2週の平日の累計売上個数」、「第2週の土日の累計売上個数」、「商品規格」、「売価」

とすること。

### 3.2 コンビニ会社からのコメント

前項三項目の打診に対するコンビニ会社からのコメントは、以下の通りであった。

- (打診1に対して): 賛成である。
- (打診2に対して): 賛成である。
- (打診3に対して): 第1週の売上動向から第2週の動向が予測できると有意義である。第4週の動向は、第2週から予測できる。コンビニ会社にとって通常、的確に予測したいニーズがあるのは第2週の動向である。

上記ヒアリングに基づき、本稿では実業務で意味のある目的変数として「第2週目の累計売上個数」を採用する。また、説明変数としては、第1週の売上動向を表現するものを採用することを目指し、次節で具体的にどのような変数が望ましいかの検討を行なう。

## 4. 目的変数を固定しての説明変数の導出

A社提供のPOSデータを分析して説明変数の導出を行なった。

### 4.1 データ概要

今回使用したデータの概要は以下の通りである：

- 中心市街地1店、郊外2店の直営店計3店舗分のデータ。
- 商品カテゴリは文献[3]と同様、「カップラーメン」と「おにぎり」の二つ。その理由は、双方ともコンビニの主力商品であることと[12]、商品の消費期限が比較的長めのものと短めのものを用意するため。
- 期間と商品数： 期間は2003年9月1日から2004年2月29日までの6か月間。この期間にカップラーメンで71商品、おにぎり

で46商品が新発売されている。

- データ項目： JANコード、商品名、商品規格、過去4週それぞれの平均売上数、曜日ごとの売上数、売価、メーカー名。

### 4.2 グラフ化による概略理解

発売初日からの経過日数と累計売上個数との相関をグラフ化して分析する(図1, 2)。3店舗共通の性質として

- おにぎり、カップラーメン共に、新商品投入から4, 5日間の累計売上個数で「売れる商品」「売れない商品」の傾向が見えてくる。
- 発売開始からの数日間の累計売上個数の傾きが急なものほど、発売2週目もそのまま順調に、累計売上個数が多くなる傾向にある。

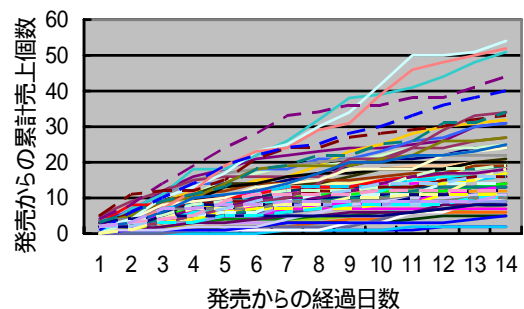


図1: カップラーメンの新商品発売後の累計売上個数の推移。

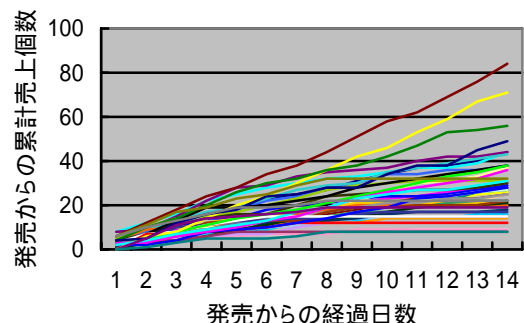
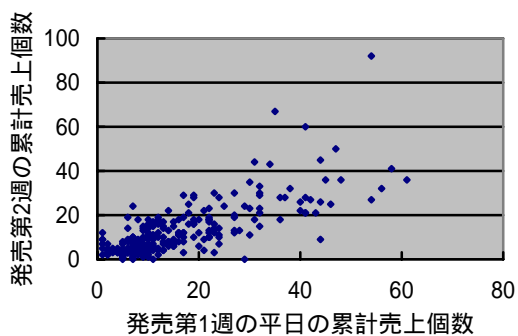
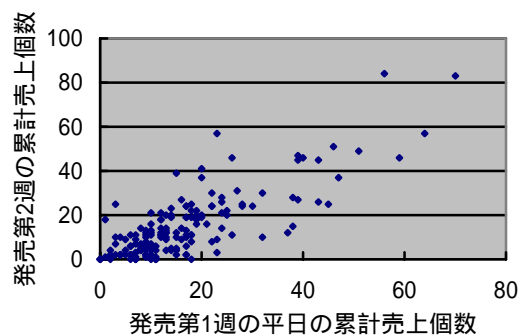


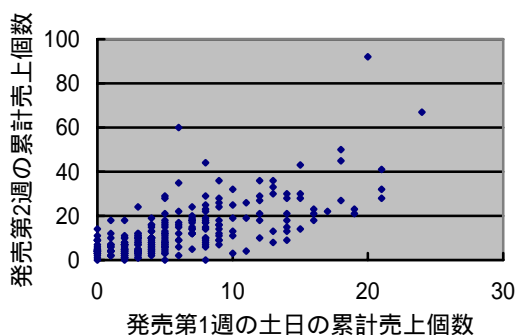
図2: おにぎりの新商品発売後の累計売上個数の推移。



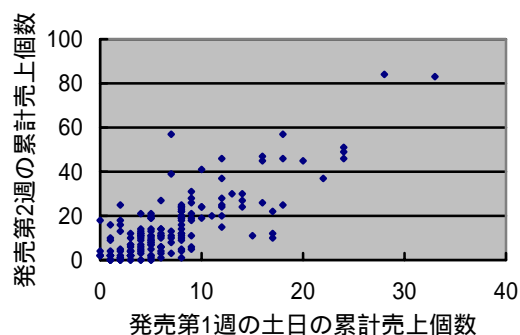
(a) 説明変数：発売第 1 週の平日の累計売上個数.



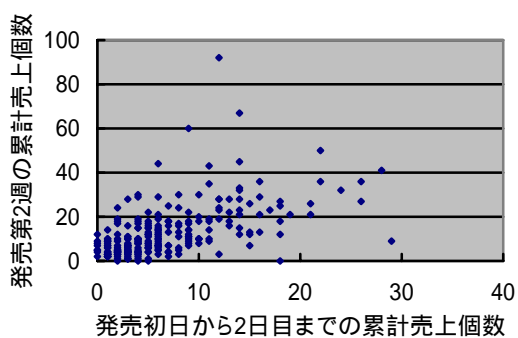
(a) 説明変数：発売第 1 週の平日の累計売上個数.



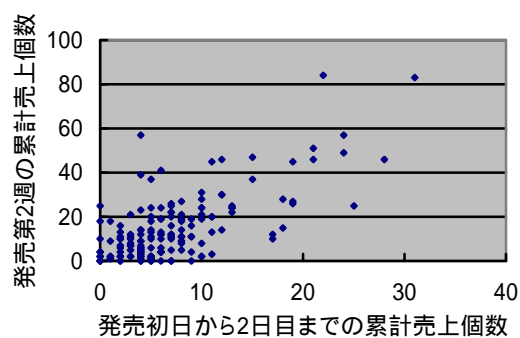
(b) 説明変数：発売第 1 週の土日の累計売上個数.



(b) 説明変数：発売第 1 週の土日の累計売上個数.



(c) 説明変数：発売初日から 2 日目までの累計売上個数.



(c) 説明変数：発売初日から 2 日目までの累計売上個数.

図 3 目的変数「発売第 2 週の累計売上個数」に対する各種変数の散布図[カップラーメンの場合].

図 4 目的変数「発売第 2 週の累計売上個数」に対する各種変数の散布図[おにぎりの場合].

### 4.3 散布図と相関係数からの説明変数の導出

目的変数を「第2週目の累計売上個数」に固定し、散布図の作成と相関係数の導出を行なった。紙幅の都合により、散布図については図3、図4に、相関係数については表1に、代表的な結果のみ記載する。

表1：説明変数ごとの相関係数

説明変数	相関係数	
	カップ ラーメン	おにぎり
投入1日目の累計売上個数	0.4760	0.6064
投入2日後の累計売上個数	0.5598	0.6918
投入3日後の累計売上個数	0.6710	0.7151
投入4日後の累計売上個数	0.7222	0.7684
投入5日後の累計売上個数	0.7508	0.7862
投入6日後の累計売上個数	0.7642	0.8086
投入7日後の累計売上個数	0.7645	0.8248
第1週の平日の累計売上個数	0.7508	0.8050
第1週の土日の累計売上個数	0.6959	0.8020
価格	-0.1545	0.0357
規格	-0.1920	

表1が示すように、

**(知見1):** カップラーメン、おにぎり共に、累計売上個数は投入後の日数が経過するごとに、相関係数がほぼ単調に増加する傾向にある。

**(知見2):** カップラーメン、おにぎり共に、日数は同じ2日間でも、投入2日後の累計売上個数より、土日2日間の累計売上個数の方が、相関係数が大きい。

**(知見3):** 新商品の投入日が月曜日以外の場合、「投入5日後の累計売上個数」と「第1週の平日5日間の累計売上個数」は異なる期間の数値であるが、おにぎりについては、

前者よりも後者の方が相関係数が大きい。

これら三つの知見は、一般性という点において決して低くない知見と見なせる。なぜならば、文献[3]で用いた市販データの分析結果(表2)からも同様の知見が得られ、相異なるコンビニ会社において共通に得られた知見であるからである。

表2：説明変数ごとの相関係数

(参考：文献[3]のデータの場合)

説明変数	相関係数	
	カップ ラーメン	おにぎり
投入1日目の累計売上個数	0.3411	0.6874
投入2日後の累計売上個数	0.5701	0.7861
投入3日後の累計売上個数	0.5733	0.7886
投入4日後の累計売上個数	0.6670	0.7854
投入5日後の累計売上個数	0.6656	0.7921
投入6日後の累計売上個数	0.7504	0.8109
投入7日後の累計売上個数	0.6956	0.8135
第1週の平日の累計売上個数	0.5288	0.7959
第1週の土日の累計売上個数	0.6546	0.8188
価格	0.0846	-0.4816
規格	0.0347	

## 5 店舗単位でのルール抽出と評価

前節で、ある目的変数とそれに対する説明変数の一般的な傾向を得たが、ここから全コンビニ会社の全店舗に数値を含めて適用できる汎用の予測式を作ろうとするのは、適切とはいえない。コンビニでは立地条件その他、店舗の特性にバラつきが少なくないからである。そのため1節で紹介した(問題点2)が生じるのであり、実際に店舗単位で決定木作成や相関ルールの抽出を行なうと、店舗ごとにバラつきの少なくない結果が得られる。(問題点1)への解決策として、一

般化は目的変数と説明変数の導出までで留めるのが適切といえる。

### 5.1 ルール抽出

前節で採用した目的変数について、「価格」や「規格」といった基からあるデータに加えて、表 1, 2 中で相関係数が上位に位置している変数群を説明変数として含めて、店舗ごとに決定木を C5.0[17]で作成した。そしてサポート値、コンフィデンス値[18]を考慮の上、合計 27 の相関ルールを抽出した。図 5 はそれらの例である。

- ・(ルール 1)： U 支店のカップラーメンは「発売 6 日目までの累計売上個数が 19 個以下」かつ「売価が 120 円以下」の場合、「第 2 週の売上個数は 0 から 6 個」である。
- ・(ルール 2)： V 支店のカップラーメンは「発売 5 日目までの累計売上個数が 15 個より大きく」かつ「発売曜日が火曜日の場合」、「第 2 週の売上個数は 21 個以上」である。
- ・(ルール 3)： W 支店のおにぎりは「平日の売上個数が 11 個以下」かつ「売価が 130 円以下の場合」の場合、「第 2 週の売上個数は 0 から 6 個」である。

...

### 5.2 抽出されたルールの評価

A 社へこの 27 ルールを報告し、これらのルールが実業務に有用なものであるか否かの評価を依頼した。具体的には、各ルールに対して「面白い」、「常識的事実」、「妥当性に疑問」、「評価不能」の 4 段階で評価の付与を依頼した。その結果、17 ルールで「面白い」、7 ルールで「常識

的事実」、3 ルールで「評価不能」という評価を得た。以上より、抽出されたルールは、ほぼ有用であったといえる。

その他に得られたコメントを整理すると、以下のとおりである。

- 導出されるルールの前提の部分が 2 つから 3 つくらいであると、解りやすく実践に使えるルールといえる。
- 逆に 4 つ以上になると、パッと見て理解出来ないもので、即実用とは行かない。
- 販売戦略がルールに反映されているものがある。
- コンフィデンス値は 7 割程度あれば、そのルールを用いた戦略にチャレンジする価値がある。

## 6. 結論と今後の展望

コンビニにおける新商品の売れ行き予測を目的として、新商品投入直後の売れ行き動向を支配するルールにおいて有用となる説明変数と目的変数の導出を行なった。

店舗を問わず有効な目的変数の一つとして「新商品投入後 2 週目の週単位の累計売上個数」をヒアリングにより得た。これに対する説明変数としての有効度において上位に位置すると推測されたのは、「新商品投入後 1 週目の土日の累計売上個数」、「新商品投入後 n 日後の累計売上個数（ただし n=7, 6, 5, 4 等）」であった。「新商品投入後 1 週目の平日の累計売上個数」については、おにぎりで説明変数として有効との結果が出たが、カップラーメンについては明確な結論を得るには至らなかった。これらの説明変数を含めて店舗単位で得た相関ルールが、実務上も有効との評価を得た。

今後の展望として、(i) 業務上有益という条件を満たす、有効な目的変数と説明変数の組み合

わせを他にも導出すること, (ii) 「おにぎりとか  
ップラメン」以外の商品カテゴリについての検  
討等が考えられる。

## 謝辞

本研究を行なうに当たって POS データを提供  
してくださった東北スーパー株式会社様に謝意を  
表します。

## 参考文献

- [1] 鷲尾隆：「ビジネスにおけるデータマイニングの現在・未来」, 情報処理学会誌情報処理, Vol.42, No.5, pp.467-471, 2001.
- [2] Private Communications, 大阪大学産業科学研究所 鷲尾隆教授, 人工知能学会第 19 回 AI シンポジウム「ビジネスプロセスに繋がるデータマイニング - 概念と実践 -」にて, '04-3-29.
- [3] 菅野憲明, 高山毅, 池田哲夫：「コンビニエンス・ストアにおける新商品発売時のデータマイニングを用いた売れ行き予測」, 情報処理学会第 130 回データベースシステム研究会/第 71 回情報学基礎研究会 合同研究発表会研究報告, DBS-130-17, FI-71-17, 2003.
- [4] 日経 QUICK 情報株式会社 (現, 日経メディアマーケティング株式会社) : NEEDS-SCAN/CVS レシートデータ, [http://www.nqi.co.jp/service/pos/cvs\\_a.html](http://www.nqi.co.jp/service/pos/cvs_a.html).
- [5] Private Communications, 慶應義塾大学環境情報学部 清木康教授, '03-5-23.
- [6] 津本周作, 平野章二, 津本優子：「医療業務基幹系へのデータマイニングの適用」, 人工知能学会研究会資料, SIG-J-A302-06, pp.31-36, 2004.
- [7] 日経リサーチ：「POS データに対するデータマイニング事例集」, 日経リサーチ販売有料資料(2000).
- [8] 「売れ行き予測システム 天気で知る来店者数 在庫調整に威力発揮」, 日経ビジネス, pp.66-68, 1996 年 9 月 9 日号.
- [9] 「IBM INFOREM のご紹介」 : <http://www-6.ibm.com/jp/dist/s3-03.htm>, (2003 年 2 月 12 日アクセス確認).
- [10] 「マーケティングリサーチ用語 60」, <http://www.intage.co.jp/report/word60/121.html>
- [11] Private Communications, 明治学院大学経済学部教授・清水聰氏：情報処理学会第 65 回全国大会・チュートリアル(2) ビジネス分野でのデータマイニング「ビジネス分野での実際の事例と応用」にて, '03-3-25, 於・東京工科大学.
- [12] 「おにぎり」, 「寿司」投入数と売れ行きの不均衡」, 業界誌コンビニ. pp.29-31, 1999 年 6 月.
- [13] 国立情報学研究所 : NACSIS-IR, <http://www.nii.ac.jp/ir/>
- [14] 情報処理学会 : 電子図書館, <http://www.bookpark.ne.jp/ipsj/>
- [15] 電子情報通信学会 : Transactions Online, <http://search.ieice.org/>
- [16] 人工知能学会 : オンラインジャーナル提供システム, <http://tjsai.jstage.jst.go.jp/ja/>
- [17] マイケル J・A・ベリー, ゴードン・リノフ著, SAS インスティテュートジャパン, 江原淳, 佐藤栄作共訳：「データマイニング手法」, 海文堂(1999).
- [18] 福田剛志, 森本康彦, 松澤裕史：データベース最前線第 10 回「データマイニング - 巨大データからの知識発見(1)」, bit, Vol.31, No.10, pp.77-83, 2000.