

インターネット CRM におけるデータマイニングの位置づけ

金子 武 久

目 次

1. はじめに
 2. リレーションシップ・マーケティングと CRM
 3. データマイニング
 4. インターネットを通じた CRM
 5. インターネット CRM におけるデータマイニングの位置づけ
 6. 顧客ユーザーのための情報処理支援システムの必要性
- 参考文献

1. はじめに

学者や研究者の研究対象でありコミュニケーションツールであったインターネットが1995年のWindows95の発売を契機にして一般の人々へと爆発的な普及をみせた。その後、インターネットの商利用が急激に増加し、インターネット上のWebサイトとしてたくさんの店舗が開設されてきた。当初は個々のサイトで独立して開設されてきた店舗であったが、顧客ユーザーのアクセス容易性を狙って、個々の店舗を一まとめにするショッピングセンター的なサイトが立ち上がってきた。このようなサイトをポータルサイトと呼ぶことも多い。2004年現在では、楽天 ichiba (<http://www.rakuten.co.jp>) や Yahoo! (<http://www.yahoo.co.jp/>) はそのなかでももっとも有名なサイトであろう。

本研究ではインターネット上の個々の店舗に焦点をあて、その店舗が経営をうまく行うためにデータマイニングがどれほど役に立ちうるのかについて考察しその可能性と限界を見極めたい。

2. リレーションシップ・マーケティングと CRM

近年、製造業者や小売業者は顧客を単なる商品の買い手とはみなさなくなっている。1990年代のバブル崩壊以後、消費者は商品を買わなくなったと言われ、売り手である製造業者や小売業者は買い手である顧客に商品を売りっぱなしにするのではなく、売ったあとでも顧客との繋がりを維持し、商品の調子や具合について尋ね、商品の故障や不具合などに積極的に取り組むようになってきた。これらは売り手が買い手を一回だけの取引相手とみるのではなく、引き続き二回

目、三回目、さらには買い手を生涯全般にわたる長期的な取引相手とみなすようになったことを表している。消費者などの顧客はもはや単なる買い手ではなく、次なるビジネスのためのアイデア創出のパートナーという考え方が広まってきた。この考え方を強調したマーケティングをしばしばリレーションシップ・マーケティングと呼ぶことが多い。

顧客との長期的なパートナーシップを形成・維持するためには、顧客のことをよく知り、その上で一人一人の顧客に対して適切な対応を行うことが必要になる。このように、一つの企業が一人の顧客に対して特別な対応を行うマーケティングを one-to-one マーケティングと呼ぶ。これを行うには、一人の顧客についてできるだけ多くのデータを収集し、それをいかに素早くそしてうまく解析・検討し、その結果得られた知見にもとづいて顧客への働きかけをすることが重要となってくる。筆者は one-to-one マーケティングは日本の伝統的な商売方法である「御用聞き」と本質的にほとんど同じであると考えている。ここで「ほとんど」と曖昧な表現をしたのは、御用聞きがごく狭い地域を商圏としていたのに対して、one-to-one マーケティングは広い地域、国、世界をも商圏としている点が異なっていると考えたからである。御用聞きは、顧客の家族構成、その嗜好、および冠婚葬祭の時期など、顧客についてのさまざまな情報にもとづいて顧客に成り代わって考えることで欲求の先読みをする。その先読みした欲求に対して適切な商品を適切な時期に推奨するということが御用聞きでは行われる。いわば御用聞きは顧客の情報処理支援を行っていることと解釈することも可能であろう。

御用聞きで商売を行っていた酒屋や米屋は商圏が狭いこともあり、近隣の大型スーパーやディスカウントショップ等に顧客を奪われ、近年その数を減少させてきているが、その商売のやり方は、顧客の情報処理支援という意味で、これからの企業 Web サイトのリレーションシップ・マーケティングを狙った運営にも大いに参考になると期待している。

顧客データには、顧客一人一人の氏名、住所、年齢、家族構成などの人口統計的な項目にはじまり、購入履歴、購入時点の状況、などが含まれる。これらの顧客データはたくさんの顧客のたくさんの購入結果をデータとして収集・蓄積しているのでかなりの大規模なデータとなっていることが多い。一般に、大規模なデータをさまざまな視点で解析し、そこからなんらかの有意義な情報を引き出すことをデータマイニングと呼ぶ。一見、無意味に思える数字や文字などのデータの山の中から、埋もれた有意義な情報を発掘するというのがこの呼び名の由来である。

したがって、リレーションシップ・マーケティングにおいては、大規模な顧客データをマイニングすることで得られた意味ある情報を活用し、顧客とのパートナーシップを形成しようとするのがしばしば行われる。このようなことを CRM (Customer Relationship Management: 顧客関係性管理) と呼ぶことも多い。リレーションシップ・マーケティングと CRM という言葉はその境界線がかなり曖昧なまま一般に用いられている。そこで、この論文では、CRM をデータマイニングなどの解析を踏まえたリレーションシップ・マーケティングという意味で用いることにする。CRM の課題としては、データマイニングなどの解析を踏まえた上で、(1)顧客とのつながりをどのようにして生み出すか、(2)顧客とのつながりを長期的にどのようにして維持するか、を

あげることができよう。この課題をクリアしつつ、顧客による一時的な買上増ではなく、長期的視点に立った顧客の一生涯にわたる買上増を狙っている。

さて、現実の製造業者や小売業者が one-to-one マーケティングを実施することには大変な困難が伴う。特に顧客が一般消費者である場合、テレビ・ラジオ、郵便、新聞・雑誌などでは顧客と十分なコミュニケーションをとることが難しく、売り手と買い手との間に明確なコミュニケーションチャンネルが存在していないからである。せいぜい、ダイレクトメールなどの即時性と双方向性、および情報量に欠けるチャンネルがあったり、主に音声のみによるコミュニケーションで情報量に欠ける電話があったりする程度であった。

3. データマイニング

3-1. データマイニングの目的

データマイニングの目的としてはレベルの異なる次の2つをあげることができよう。

- 大量データに潜む法則性や傾向、パターンの発見
- パターンをビジネス等の利用目的に照らして選別・解釈することによって有意味化

一般的にはデータマイニングの目的と言えば前者を指す。しかしながら、今日のような POS データや Web ページアクセスログデータなどはさわめて膨大であり、そのデータを解析することによってたくさんの統計的な法則性やパターンが発見されるようになった。いわば、法則やパターンの洪水に直面し、どれが役立つ法則・パターンなのかが問題となり、利用目的に沿って新たに法則やパターンの絞り込みをする必要が出てきたのである。そこで第2の目的が浮上してきた。

ここで例として有名な「缶ビールと紙おむつ」を考えてみよう。データマイニングの一手法であるアソシエーション分析に POS データをかけることでたくさんの同時購入パターンが発見された。その一つのパターンが「缶ビールと紙おむつ」の同時購入である。普通、このパターンの解釈は次のようになされている。缶ビールと紙おむつが同時に購入されることが多いのは、これらを購入しているのが若い夫婦と赤ちゃんによる世帯で、彼らは若いため収入が多くなく、まして育児に費用がかかるためなかなか贅沢な生活ができない。したがって、育児のために紙おむつを買い、アルコール類についてはブランデーやワインなどの高額なものではなく安価な缶ビールを買い求めるのだ、ということである。このような解釈のもとで、例えば、赤ちゃんのいる若い世帯には生活のための低価格品をアピールすればよい、というようなビジネスに結びつけられた仮説的法則が導かれてくる。ここに至って、ようやく、データマイニングの利用目的に照らして意味のあるパターンが発見できたことになる。「缶ビールと紙おむつ」の同時購入が多いというのは確かに一つのパターンではあるが、それだけでは無意味なパターンである。特に、マーケティングや経営という分野でデータマイニングを利用する場合はなおさらである。

ところで、上のような解釈の他に、次のような解釈も可能であろう。缶ビールと紙おむつが同時に購入されることが多いのであれば、スーパーの店内の売場配置において、缶ビールと紙おむ

つを遠くに離して陳列しておけばよい。そうすれば、缶ビールを買う人は缶ビールをショッピングカートに入れた後、紙おむつを買うために長い距離を移動することになり、その分、多くの商品に接する機会を得て、衝動買いを促すことになる。顧客の一人当たり買上単価を増やすためには同時購入される商品同士を遠く離して陳列すればよいことになる。

さらに一方で次のような解釈も可能であろう。缶ビールと紙おむつが同時に購入されることが多いのであれば、顧客の買物時の利便性を考えて両者を隣同士に陳列する、ということである。こうすれば、顧客は缶ビールの後、続けて紙おむつをショッピングカートに入れることができ買物がしやすくなり、お店への愛顧心が向上し、買上の増加をもたらしてくれることになる。

データマイニングの一般的な目的は大規模データに潜む法則性やパターンを発見することであるが、上で検討したように、それだけでは無意味な法則性やパターンであることが多い。その法則性やパターンを有意味化するのが解釈にはかならない。しかしながら、その解釈は一通りには決まらないことが多い。上で見たように、缶ビールと紙おむつの同時購入という一つのパターンに焦点をあてても、3つもの解釈が出てきた。もっと考えれば4つ以上の解釈をすることも可能であろう。発見されたパターンの利用目的に照らし合わせた解釈まで含めてデータマイニングの目的としなければ有意味な発見には繋がらないと言える。

上では一つの統計的なパターンから3つの解釈を生み出した。通常、それらの解釈は新たなパターン発見へ向けての仮説を構成することが多い。例えば、上では、缶ビールと紙おむつの同時購入が多いという統計的なパターンにもとづいて、両商品を店内で遠くに離して陳列すればよいという実用的な仮説パターンと、両商品を近くに陳列すればよいというもう一つの実用的な仮説パターンを生み出した。この二つの仮説は互いに相反するもので、この段階ではどちらがもっともらしいかを判定することは困難である。そこで、店頭実験を用意し、遠くに陳列した場合と近くに陳列した場合の売上に対する貢献度を調べることになる。得られたデータを解析することによってどちらの仮説がもっともらしいかを統計的に検定することが可能となろう。しかしながら、これは売上の増大という視点での比較研究であり、顧客の買物利便性の向上という視点での比較研究とは異なる結果をもたらすことになろう。

以上によって、データマイニングはただのパターンを発見するための道具や作業ではなく、利用目的を考慮に入れた有意味なパターンを発見することが重要であることが理解できよう。

3-2. データマイニング技法とパターン識別

一口にパターンと言ってもさまざまなものが考えられる。あるものとまたあるものがどのような関係になっているかについてのパターン、たくさんのものがどのように分類・要約されるかについてのパターンなどがある。ここでは前者を関係性パターン、後者を分類パターンと呼ぶことにする。

データマイニングの技法には実に多くのものが存在する。従来、多変量解析法と呼ばれていた一群の解析技法を含め、最近になって発展してきたニューラルネット、ベイジアンネットなどさ

さまざまな技法があげられる。ここで各種のデータマイニング技法を関係性パターンを把握するのによく用いられるものと、分類パターンを把握するのによく用いられるものに分類してみた。

• 関係性パターン識別
回帰分析, 判別分析, 正準相関分析, 数量化 I 類, 数量化 II 類, コンジョイント分析
ニューラルネット
隠れマルコフモデル, N グラムモデル, 最大エントロピーモデル
時系列モデル
リンク分析, アソシエーション分析, ペイジアンネット
• 分類パターン識別
決定木
協調フィルタリング, コンテンツベースフィルタリング
主成分分析, 因子分析, 正準相関分析, 数量化 III 類, 数量化 IV 類
多次元尺度法, 階層型クラスタ分析
非階層型クラスタ分析
自己組織化マップ
ABC 分析

これらのデータマイニング技法はさまざまな研究分野で誕生し、発展してきたものである。例えば、ニューラルネットは脳神経学、N グラムモデルや最大エントロピーモデルは確率言語学、隠れマルコフモデルは遺伝学における DNA 解析や確率言語学、時系列モデルは計量経済学、決定木や協調フィルタリングやコンテンツベースフィルタリングは人工知能、などである。

ここでの技法の分類は客観的で確定的なものではなく、主観的でやや恣意的なものである。関係性パターンと分類パターンは相互に繋がりがあがり、きれいに二分できるものではない。とは言うものの、ここでの分類は大方のデータ解析の関係者にとっては了解のできるものであろう。

4. インターネットを通じた CRM

4-1. インターネットの特徴

近年のインターネットビジネスの成長は目を見張るものがある。インターネットのメディアとしての特徴をあげるとすると次のようになろう。

- 双方向性
- 即時性
- 広範囲性
- マルチメディア情報の取り扱い

双方向性は、情報が行ったり来たりすることで情報の送り手と受け手が相互に交代可能であることを指す。これまで電話は双方向型のコミュニケーションメディアの代表であった。テレビやラジオといったメディアはあまり双方向型ではなく、どちらかというと一方向性の強いコミュニ

ケーションメディアである。インターネットは電子メールや Web ページでの掲示板システムやチャットシステムなどを通じて双方向型のコミュニケーションをすることが可能である。

即時性は、情報の送り手が情報の追加・更新をした場合、受け手は即座にその追加変更を知ることができることを指す。例えば、電話、テレビ、ラジオ、新聞は即時性が高いコミュニケーションメディアと言えるが、雑誌は即時性が低いと言える。インターネットでは即時性の高いコミュニケーションが可能である。

広範囲性とは、情報の伝達範囲の広さを指しており、電話や手紙などは伝達範囲が極めて狭いコミュニケーションメディアである。一方、テレビやラジオ、雑誌や新聞などは比較的伝達範囲が広いコミュニケーションメディアと言える。インターネットは、例えば一つのサイトの Web ページが全世界のインターネットで結ばれたコンピュータへと情報伝達が可能であるからきわめて広範囲性の高いコミュニケーションメディアと言うことができる。

マルチメディア情報の取り扱いとは、さまざまなタイプの情報がやりとりされることを指す。例えば、手紙では文字情報が中心、ラジオでは音声情報のみ、テレビでは音声、映像、文字情報など複数のタイプの情報がやりとりされる。インターネットではテレビと同じようにさまざまなタイプの情報がやりとりされ、マルチメディア情報を扱うことができる。

以上のように、インターネットは双方向型で、即時性が高く、広範囲にマルチメディア情報をやりとりできるコミュニケーションメディアであると言うことができ、他のコミュニケーションメディアと比較した場合、もっとも現実的なコミュニケーションに近いものである。なお、ここでいう現実的なコミュニケーションとは二人の人間が相対して情報伝達することを指す。現実的なコミュニケーションは人間のあらゆる感覚器に訴えかける情報伝達が可能であり、そこでは音声、映像・文字、接触、香り、味覚などのマルチメディア情報をやりとりすることができる。

4-2. インターネット CRM

このような特徴をもつインターネット上でたくさんの店舗が開設され、それぞれの店舗では多種多様な商品が販売されているに至っている。インターネット上で開設されている店舗や企業が運営するサイトを以下では企業 Web サイトと呼ぶことにする。企業 Web サイトにおいても顧客との繋がりを重視し、リレーションシップ・マーケティングを展開しているところが現れてきている。企業 Web サイトはインターネットを通じて特定の顧客とのコミュニケーションチャネルをもつことができる。これと先にあげたインターネットの特徴、すなわち、双方向性、即時性、広範囲性、マルチメディア情報のやりとりとが相まって、企業 Web サイトは顧客とのこれまでになかった密度の濃いコミュニケーションをすることが可能になってきたのである。

企業 Web サイトがリレーションシップ・マーケティングを展開する際、重要となってくるものは自らの企業 Web サイトの情報とそこにアクセスしてくる顧客の情報である。これらの企業 Web サイト情報と顧客情報をデータマイニングすることによって明らかにされる何らかの法則性やパターンにもとづき、顧客との結びつきを強化すべく企業 Web サイトの運営に役立ててい

こうとすることを「インターネット CRM」と呼ぶことにする。

インターネット CRM をするにあたって、どのような企業 Web サイト情報と顧客情報が存在するのであろうか。ここでこれらについて若干の整理をしておきたい。

企業 Web サイトに関連する情報としては次のようなものがあげられよう。

〈企業 Web サイト情報〉

- 企業 Web サイトを構成しているページ属性
- ページ間のリンク構造
- 企業 Web サイト管理者によるコンテンツ（記事・コメント、評価ポイント、静止画、動画、音声など）
- 顧客によるコンテンツ（記事・コメント、評価ポイント）

ここで、「企業 Web サイトを構成しているページ属性」とは、当該ページに記述されている文字数、画像データの大きさ、画像データの数、他ページへのリンク数、標準的なブラウザの窓サイズにおける一画面中の文字情報の比率や画像情報の比率、などが含まれる。また、「ページ間のリンク構造」とは当該サイト全体のページがどのようにリンクづけされているかを表したもので、主として隣接行列として表現されるものである。

続いて顧客に関する情報としては次のようなものがあげられよう。

〈顧客情報〉

- Web アクセスログ
- 購入履歴
- 顧客ユーザー属性

ここで、「Web アクセスログ」は Web サイトのサーバー上に蓄積される顧客ユーザーのページアクセス履歴を指す。「購入履歴」とは Web アクセスログの一部ともみなせるものであるが、企業 Web サイト上での顧客ユーザーによる商品・サービス等の購入を記録したデータを指す。「顧客ユーザー属性」とは、アクセスしてきた顧客ユーザーの氏名、年齢、性別、職業、住所、電話番号、クレジットカード番号などの個人情報はもちろん、顧客ユーザーの性格（例えば、革新的⇔保守的、落ち着きのある⇔せっかちな、など）といった情報も含まれる。

以上のような企業 Web サイト情報と顧客情報についてデータマイニングを行うことによってビジネスに有効なパターンを発見し、そのパターンをうまく利用することによって企業 Web サイトと顧客との関係性を強めていこうというのがインターネット CRM の目的と言えよう。

5. インターネット CRM におけるデータマイニングの位置づけ

5-1. インターネットを通じたりレーションシップ・マーケティングの課題とその解決手段

企業 Web サイトへの顧客ユーザーのアクセスと関連してインターネットを通じたりレーションシップ・マーケティングを行うための具体的な課題とその解決手段を整理してみた。企業 Web サイトと顧客との関係性を強くするためには、(1)多くの新規顧客ユーザーを獲得すること、

(2)新規顧客ユーザーをリピータへと育てること、(3)既存顧客ユーザーによるリピータアクセスを維持・増加させること、の3点が重要と考えられる。それぞれの課題についてどのような解決手段があるのかを検討し、それぞれ次のようなものを考えた。

(1) 新規顧客ユーザーの獲得

- 有名ポータルサイトでのリンクづけ
- 検索エンジン向けの適切なキーワード使用
- 実店舗との連携（チラシ配り、キャンペーン、懸賞、etc.）
- 顧客ユーザーのコミュニティ形成支援（掲示板提供、etc.）

(2) 新規顧客ユーザーのリピータ化

- 新たな顧客ユーザー欲求の把握とサービスの拡大
- サイト内のリンク構造の明確化

(3) 既存顧客ユーザーのリピータを維持・増加

- コンテンツの充実化
- Web サイト内ページングの高速化
- Web ページのデザイン工夫（色組合せ、素材配置、etc.）
- 顧客ユーザーのコミュニティ支援
- 推奨プロモーションなどの高度化

以上のような課題に対して解決手段を適切に運用・改善することによって、企業 Web サイトと顧客ユーザーとの関係性は強化されていくものと考えられる。さて、このような解決手段を適切に運用・改善するために、企業 Web サイト情報や顧客情報をデータマイニングすることがどのような役割を果たすのであろうか。本論文ではインターネット上でデータマイニングを通じてリレーションシップ・マーケティングを展開することをインターネット CRM と呼んでいたが、この疑問に答えることはインターネット CRM の存在意義を確認するために重要である。

5-2. インターネット CRM の可能性

インターネットを通じてリレーションシップ・マーケティングを展開するための具体的な課題とその解決手段について上で記述した。この解決手段を適切に運用・改善するために、データマイニングは役に立つのであろうか。以下ではこの点について考察しておこう。

5-2-1. 新規顧客ユーザーの獲得のためのデータマイニング

新規顧客ユーザーを獲得するためには、当該サイトの内部はもちろん、それ以上に、サイトの外での工夫が重要となる。例えば、新規顧客ユーザーの獲得のために、有名ポータルサイトでのリンクづけ、検索エンジン向けの適切なキーワード使用、実店舗との連携（チラシ配り、キャンペーン、懸賞など）、顧客ユーザーのコミュニティ形成支援（掲示板提供など）などの解決手段が考えられた。有名ポータルサイトでのリンクづけをするのにデータマイニングが役に立つのであろうか。あるいは、検索エンジン向けに適切なキーワードを使用するためにデータマイニング

が役に立つのであろうか。また、実店舗との連携をするためにデータマイニングが役に立つのであろうか。顧客ユーザーのコミュニティ形成支援をするためにデータマイニングが役に立つのであろうか。

ポータルサイトへのリンクづけは、データマイニングによらなくてもしかるべき広告料を支払うことで登録すればよいことである。これによって新規顧客ユーザーの獲得に大いに貢献してくれよう。

続いて、検索エンジン向けに適切なキーワードを使用することについても、広告料を支払うことで検索エンジンにしかるべきキーワードを登録すればよい。しかしながら、検索エンジンでなんらかのキーワードを入力してから当該サイトに訪れてきた顧客ユーザーが、サイト内の1~2ページを短時間で閲覧してすぐに他のサイトへと離れて行ってしまった場合、顧客ユーザーが用いたキーワードの意図と検索エンジンで検索されてきたサイトとが合致していなかったことが考えられる。このような顧客ユーザーがどの程度存在するのかはアクセスログデータを解析することで把握することが可能である。しかるべき設定をすれば、現在閲覧しているページの直前にどのページを閲覧していたのかがシステム的环境変数\$HTTP_REFERERに格納されているので、これをアクセスログデータに残しておくことができる。これを利用して解析すればよい。もし多数の顧客ユーザーがこのような見当違いの検索結果にもとづいて当該サイトに訪れてきているとしても、このような顧客ユーザーを排除すべきであるという理由は存在しない。こういった顧客ユーザーをいかにしてサイト内に滞在させるかがサイト設計者の腕の見せどころとなろう。一方、顧客ユーザーがキーワード検索を行っても当該サイトに訪れることができなかったという場合もある。顧客ユーザーの意図と使用したキーワード、サイトの意図・目的とそこの各ページで使われているキーワード、これらの適切なマッチングが重要である。顧客ユーザーの意図とサイトの意図・目的は合致しているのに、それぞれの使っているキーワードが不適切であったためにきちんと検索されてこないということが大いに考えられるからである。しかしながら、これらをデータマイニングで明らかにすることは現在のところ不可能であると言わざるを得ない。

実店舗との連携をすることについては、例えば、実店舗で販売されている商品に何らかの識別番号を印刷しておき、その識別番号を手がかりにしてしかるべきサイトにアクセスすると懸賞に応募できるような仕組みが稼働している。インターネットをプロモーション手段として利用するこのようなやり方がどの程度の効果があるのかをデータマイニングで探ろうというわけである。しかしながら、これはこれまで行われてきた懸賞がインターネットを利用するようになっただけのことであり、なんらかの購入促進をもたらすことは十分予想できる。わざわざ手間をかけてデータマイニングを行う意味がどこにあるのだろうか。また、インターネット上で売れている商品が実店舗でも売れるのだろうか。あるいはまた、その逆は有り得るのか。これらを調べるためにデータマイニングはなんらかの役に立ちそうではあるが、実際どの程度の役に立ちうるのかについてはやってみなければ分からない。

顧客ユーザーのコミュニティ形成支援をするためにデータマイニングが役に立つのであろうか。

顧客ユーザーのコミュニティと言えば、掲示板システムやメーリングリストが一般的である。このような場ではさまざまな情報が飛び交い、顧客ユーザー同士の情報交換が行われる。これらに役立つものとしては、掲示板システムに投稿された記事を整理要約するテキストマイニング、キーワードによる記事検索システムなどが考えられる。

5-2-2. 新規顧客ユーザーのリピータ化のためのデータマイニング

当該サイトに訪れた新規顧客ユーザーの欲求がこれまで当該サイトに訪れた既存リピート顧客ユーザーと同様であるなら、特に問題はないものと考えられる。ある種の欲求を充足させるべく既にリピート顧客ユーザーが存在しているなら、それと同種の欲求をもった新規顧客ユーザーもいずれそのうちにリピータになってくれることが期待できるからである。問題なのは、これまでのリピータとは異なる欲求をもつ新規顧客ユーザーをどのようにしてリピータに育て上げるかである。そのためには新たな顧客ユーザー欲求を発見し、それに向けた商品やサービスの準備が不可欠であろう。

新たな顧客ユーザーの欲求を把握するためには、数多くの試行錯誤が必要である。さまざまなアイデアの下、実験的な試みを多く実施することにより、今後伸びそうな新たな商品や商品カテゴリを見つけることができ、その商品を購入する顧客ユーザーの特徴などをデータマイニングで発見できるかもしれない。

サイト内のリンク構造が不明確である場合、顧客ユーザーは、現在、サイト内のどこを閲覧しているのか分からなくなることがある。ページ間のリンクが絡み合っている場合にそのようなことがよく起こる。いわば迷路に迷い込んだようなものであり、このようなサイトにアクセスしてきた場合、多くの顧客ユーザーはページ探索を途中で投げ出し、他のサイトへと離脱してってしまう可能性がある。また、検索エンジンから当該サイトのあるページへジャンプしてきた顧客ユーザーが、サイト内の他のページを閲覧しようとしてもそのページからリンクが貼られておらず、他ページの閲覧ができないということもある。袋小路に迷い込んだようなもので、そのページから出るに出不れないというサイトである。これでは顧客ユーザーを当該サイトにひきつけておくことは難しい。迷路にならず、袋小路にならないようにサイトのページリンクを設計する必要がある。そのために、アクセスログデータの解析が役に立つ可能性がある。アクセスログデータを解析することで、顧客ユーザーがサイトに入ってきた最初のページからどのようなページを辿ってサイトから離脱していったかが分かる。つまり、あるページを最初に訪れた場合、辿り着けないページが存在していないか、辿りつけても実際にはほとんどの顧客ユーザーが閲覧しないページが存在していないか、などを見つけることができるわけである。この結果をサイトの再設計に用いることにすればデータマイニングは役に立つと言える。

5-2-3. 既存顧客ユーザーのリピートを維持・増加

既存顧客ユーザーの当該サイトへのリピートを維持・増加させることはリレーションシップ・マーケティングにとっては特に重要な課題である。これを実現させるためには、新規コンテンツの追加や既存コンテンツの更新などによるコンテンツの充実化、Webサイト内ページングの高

速化, Web ページのデザイン工夫 (色組合せ, 素材配置, など), 顧客ユーザー・コミュニティへの支援, 推奨プロモーションなどの高度化, が考えられる。これらのためにデータマイニングはどの程度の役に立ち得るのであろうか。

コンテンツの充実化のために, 各ページに顧客ユーザーの評価を入力する欄が用意されているサイトが既に存在している。それは, 「このページはあなたの期待を満足させるものでしたか?」という問いかけに対して, 1点から5点までの評価がマウスクリックで入力できる仕組みである。このような顧客ユーザー評価ポイントを解析することによってコンテンツの充実度を把握することがある程度可能である。評価の低いページについてはそのコンテンツを更新する必要がある。この意味でデータマイニングはコンテンツの充実化に役に立ちうる。しかしながら, どのように更新したらよいのか, あるいは, どのようなコンテンツを新規追加したらよいのかはデータマイニングからは分かりそうにない。

Web サイト内ページングの高速化は顧客ユーザーの離脱を避けるために重要である。現在閲覧しているページからサイト内の他のページへ移る際, 次のページがなかなか表示されないことがある。そのような場合, 顧客ユーザーはその「待ち」が我慢できず, ブラウザに登録してあるBookmark などから他のサイトへと離脱してしまうことが考えられる。ページの表示が遅い原因としては, サーバーの処理能力が低いこと, ネットワークが混雑していること, 表示しているページに貼り付けてある画像や音声などのデータサイズが大きいこと, などが考えられる。サーバーの処理能力については, CPU の性能を上げたり, メモリを増やしたりすればよい。ネットワークの混雑については, ミラーサイトを立ち上げて複数のサイトで同一内容が見られるように工夫することが考えられる。その際に, どのようなネットワーク経路を通ってきている顧客ユーザーが多いのかを IP アドレスをもとに経路探査することによってデータを集め, これを解析することによってどのあたりにミラーサーバーを立ち上げればよいのか把握できる可能性がある。このためにデータマイニングは役に立つかもしれないが, しかしこれは1つの Web ショップが行うにはあまりに荷が重い作業であり, コスト対パフォーマンスは低いと予想される。一方, ページに貼り付けてあるデータサイズが大きい場合には, そのサイズが小さくなるような工夫をしてやればよい。この点についてもデータマイニングは使えはするであろうが, それよりにより, 自分で当該サイトにアクセスしてみれば表示の遅さが実感できるので, データマイニングの出る幕はないと言える。

Web ページのデザイン工夫 (色組合せ, 素材配置など) については, コンテンツの充実化のところで説明したように, 当該ページのデザインを顧客ユーザーに評価してもらい, そのデータを解析することでデザインの工夫に役立てることが可能である。しかしながら, どのような色の組合せや素材の配置が望ましいのかはデータマイニングからは分かりそうもない。

顧客ユーザーのコミュニティ支援については, 掲示板システムやメーリングリストなどに記載されている情報を要約したり, キーワード検索できるようにしたりすることが一般に行われている。情報要約というテキストマイニング, 検索システムといった形でのデータマイニングが役に

立っている。

また、既存顧客ユーザーを維持・増加させるために、推奨プロモーションを高度化することが考えられる。現在、amazon.co.jp などでは協調フィルタリングを用いているようである。これまで、どの顧客ユーザーがどのような商品を読覧したか、また、どのような商品を購入したかといった情報がデータベースに保存されている。このデータを解析することによってある種の類似性をもったパターンが見えてくる。例えば、A、B、Cといった3種類の商品の組み合わせをたくさん顧客ユーザーが購入していることが分かったとしよう。ここにBとCを既に購入した顧客ユーザーがいるとする。他の顧客ユーザーとの類似性を仮定するとこの顧客ユーザーは次にAを購入する可能性が高い。そこで、この顧客ユーザーに商品Aを推奨するわけである。このような推奨の仕組みをもっと高度にするようなデータマイニングの技法を開発しこれを実現することで顧客ユーザーを繋ぎとめておくことができると考えられる。

6. 顧客ユーザーのための情報処理支援システムの必要性

データマイニングの目的は先にも記したが、一言で表現すれば、役に立つパターンの発見となる。ある種のパターンを知っていれば、ビジネスの現場では、未知なる新規顧客が訪れたとき、その顧客と過去の既知の顧客との類似性に着目し、類似性が見つかったらその過去の顧客についてのパターンを利用して新規顧客に対処することができる。例えば、過去の顧客データを解析することで「20代の男性サラリーマンは青系のネクタイを購入する」というパターンが知られているとしよう。あるとき新規の顧客が来店してきた。その顧客は見たところ20代の男性でサラリーマン風である。先のパターンとの類似性が考えられるので、この新規顧客に青系のネクタイを薦めてみるという接客をすることが可能である。パターンにもとづいて何かをやるということは、顧客の特性や商品の種類などで違いはあろうが、おおよそここで記述したようなことになる。

このような顧客への接し方は顧客を「個」客と見ておらず、顧客を「集合」客と見ていることを意味する。いわば、市場細分化的なある種の同質な欲求をもった顧客集合という意味で顧客を捉えているのである。

先の章で記したように、インターネット CRM においてデータマイニングがどの程度役に立つのかについて考察してきたが、結局、「個」客対応のためのものではなく、「集合」客的に対処するのに役に立つということになる。インターネットは、双方向性、即時性、広範囲性、マルチメディア情報の取扱いを特徴としていた。世界中の顧客ユーザーと当該サイトとの間で one-to-one の対応が可能なのわけであるが、データマイニングによってパターンを発見し、これを利用することは顧客ユーザーを集めた取り扱うことになり、顧客ユーザーの「個」を見失わせる可能性がある。確かに、未知の新規顧客ユーザーが当該サイトに訪れたとき、その顧客ユーザーにどのように対処したらよいかを示すマニュアルとしての役目をデータマイニングによる成果は果たしてくれる。これは one-to-one 対応の全てではない、ということをよく承知しておく必要がある。先にも記したように one-to-one の基本は伝統的な商売方法の「御用聞き」にある。御用

聞きの本質は顧客の情報処理支援にあると言える。顧客の情報処理支援をするために、過去の顧客のパターンが役に立つこともあろう。ここにデータマイニングの存在意義があると言える。

では、過去の顧客のパターンが役に立たなくても、顧客の情報処理支援をすることは可能であろうか。そのためには、

- 商品やサービスに関するデータベースの充実化
- 検索システムの高度化
- 顧客ユーザーとのインターフェイスの高度化

を通じて、顧客ユーザーが自らの欲求を充足させるための情報処理負荷を軽減させることが可能であるとする。データベースの充実化というのは、現実性をもたせることを狙っている。商品についての一方からの写真だけでなく多方向からの写真を含めたり、実際の生活の中での当該商品の大きさをイメージさせるような工夫をしたり、さまざまなデータを加えることで現実性を帯びさせるわけである。その上で、単なるキーワード検索だけでなく、顧客ユーザーの意図や意味を汲んだ上での検索システムを用意する。さらに、顧客ユーザーと対話することによってユーザーの抱える問題や疑問などに対して適切に対応するような高度に知的なインターフェイスを採用するわけである。このように、データベース、検索システム、ユーザーインターフェイスを高度にすることは、まるで、経験豊富な販売員が企業 Web サイトに常駐しているようなものであり、これが顧客ユーザーの情報処理支援を行うわけである。現在の技術水準では企業 Web サイトの担当者が電子メールや掲示板等を通じた手作業で大部分の顧客ユーザー情報処理支援を行わざるを得ないが、技術の進歩とともにこれらの作業が自動化され、やがては企業 Web サイト上に構築された人工知能が顧客ユーザーの情報処理支援を全自動で行う時代も来ることであろう。このような流れに視点をおくと、amazon.co.jp が現在行っていることは顧客ユーザー情報処理支援の自動化に向けた大いなる予兆として期待できるものである。

以上、考察したように、インターネット CRM においてデータマイニングは顧客ユーザー対応のための「マニュアル作成」というような意味で大いに役に立ちうる事が把握された。しかしながら、インターネット CRM において one-to-one マーケティングを強調すると、マニュアル的な対応をするだけでは決して十分とは言えない。データマイニングは顧客ユーザーを集合客として扱う傾向にあるからである。御用聞きのように「個」としての顧客対応をするためには、データマイニングを含めた上で、顧客ユーザーの情報処理支援という視点で企業 Web サイトを構築・運営することが重要であろう。

参考文献

- 秋葉友良 (2002), 「自然言語処理におけるベイジアンネット」, 『人工知能学会誌』, Vol.17, No.5.
麻生英樹, 津田宏治, 村田 昇 (2003), 『パターン認識と学習の統計学計 新しい概念と手法』, 岩波書店。
有田正規 (2002), 「ベイジアンネットとバイオインフォマティクス」, 『人工知能学会誌』, Vol.17, No.5。

- 有村博紀, 浅井達哉 (2003), 「テキストマイニング: ウェブデータからの知識発見を目指して」, 『Computer Today』, No.114。
- 石井 信 (2003), 「不完全データの問題とバイオインフォマティクス」, 『Computer Today』, No.114。
- 和泉憲明, 武田英明, 山口高平 (2002), 「意味理解する Web を目指して—一次世代 Web の方向性を探る—」, 『人工知能学会誌』, Vol.17, No.4。
- 岡山将也, 小池 博 (2002), 「E-コマースサイトにおける優良顧客向けパーソナライゼーション」, 『人工知能学会誌』, Vol.17, No.5。
- 北 研二 (1999), 『確率的言語モデル』, 東京大学出版会。
- 北 研二, 津田和彦, 獅々堀正幹 (2002), 『情報検索アルゴリズム』, 共立出版。
- 金 明哲, 村上征勝, 永田昌明, 大津起夫, 山西健司 (2003), 『言語と心理の統計—ことばと行動の確率モデルによる分析—』, 岩波書店。
- キンラン J.R., 古川康一監訳 (1995), 『AI によるデータ解析』, トッパン。
- 坂和正敏, 田中雅博 (1995), 『遺伝的アルゴリズム』, 朝倉書店。
- 三宮信夫, 喜多 一, 玉置 久, 岩本貴司 (1998), 『遺伝的アルゴリズムと最適化』, 朝倉書店。
- 豊田秀樹 (2001), 『金鉱を掘り当てる統計学』, 講談社ブルーバックス。
- 馬場 肇 (1998), 『日本語全文検索システムの構築と活用』, SOFTBANK。
- 武田英明 (2002), 「知性のネットワークとしての WWW — Web インテリジェンスに関する一考察—」, 『人工知能学会誌』, Vol.17, No.3。
- 守口 剛, 阿部 誠, 金子武久, 井上達紀, 佐藤栄作 (2002), 『クリックシークエンスに影響するページ属性の把握と予測モデルの構築』, IMSI 報告書。
- 山本 学 (2002), 「E-コマースに関わるエージェント技術の動向」, 『人工知能学会誌』, Vol.17, No.5。
- Duda, Richard O., Peter E.Hart and David G.Stork (2001), *Pattern Classification*, Wiley-Interscience. (尾上守夫監訳 (2001), 『パターン識別』, 新技術コミュニケーションズ)
- Fayyad, Usama M., Gregory Piatetsky-Shapiro, Padhraic Smyth, and Ramasamy Uthurusamy eds. (1996), *Advances in Knowledge Discovery and Data Mining*, MIT Press.
- Hastie, Trevor, Robert Tibshirani and Jerome Friedman (2001), *The Elements of Statistical Learning*, Springer.
- Kargupta, Hillol and Philip Chan eds. (2000), *Advances in Distributed and Parallel Knowledge Discovery*, MIT Press.
- Kohonen, T. (2001), *Self-Organizing Maps*, Third Edition, Springer. (コホネン, T., 徳高・岸田・藤村訳 (1996), 『自己組織化マップ』, シュプリンガー・フェアラーク東京)
- Koski, Timo (2001), *Hidden Markov Models for Bioinformatics*, Kluwer Academic Publishers. Manning, Christopher D., Hinrich Schutze (1999), *Foundations of Statistical Natural Language Processing*, The MIT Press.
- Mitchell, Tom M. (1997), *Machine Learning*, McGraw-Hill.
- Srivastava, Jaideep and Robert Cooley, Mukund Deshpande, Pang-Ning Tan (2000), “Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data,” *SIGKDD Explorations*, January 2000, Volume 1, Issue 2.
- Zaiane, Osmar R., “From Resource Discovery to Knowledge Discovery on the Internet,” Ph. D. Thesis.