

●Company Profile

会 社 名	株式会社サムライファクトリー
代 表 者	代表取締役 篠田裕輔
所 在 地	東京都豊島区池袋 2-50-9 第三共立ビル
U R L	http://www.samurai-factory.jp
事 業 内 容	<ul style="list-style-type: none"> • Web パーツサービス • Web 広告サービス • マルチメディア事業 • 基幹システム構築 • Web アプリケーション開発などの法人向けサービス

**データマイニングでビジネスを
加速する！**

6

●—————●
Chapter

Web ビジネスもリアルビジネスも データマイニングで変わる

1. Web マーケティング進化の裏に データマイニングあり

◆データマイニングとは何か

データマイニングというのは、「data(データ)」を「mine(マイン)」することである。mine とは、「採掘する」とか「鉱山」という意味がある。

つまり、データマイニングとは、各種のデータを活用して、利益を生み出すようなビジネス上の金脈を採掘することを意味する。

具体的には、膨大なデータを各種の手法を使って分析し、それらの中から価値ある情報を発見することであり、そのための手法群のことをいう。

データマイニングという分析手法を実施するためには、「データマイニング・ツール」と呼ばれるものを活用するのだが、そのツールが日本に紹介されたのは、1995年頃のことであり、日本ではまだ十余年の歴史しかないということになる。

さて、データ分析ということでは、データマイニングが導入される以前から、統計解析という方法による分析は行われていたわけであるが、従来の統計解析とデータマイニングの違いを明らかにしておきたい。

そもそもデータマイニングについては、統計解析が発展した形

であるとか、統計解析とはまったく別のアプローチであるとか、実は統計解析は広い意味でのデータマイニングに内包される概念だとかいったさまざまな捉え方があった。

しかし、そうした議論は議論として、統計解析とデータマイニングのもっとも大きな違いをいえば、それは、統計解析が仮説検証のための分析手法であるのに対して、データマイニングは仮説発見のための分析手法である、という点にある。

仮説検証型というのは、文字どおり、まず仮説があって、その仮説が正しいかどうかを統計解析によって明らかにするということである。

これに対して、仮説発見型というのは、事前の仮説を必要とせず、今あるデータをすべて分析にかけて、何らかの規則性や法則性を発見してしまうというものである。

たとえば、あるスーパーマーケットで「何と何が一緒に買われることが多いか」ということを分析によって導き出そうという場合、統計解析の手法の場合には、まず「たまごとビールは一緒に買われる可能性が高いのではないか」という仮説を立て、その仮説に基づいて、「たまご」を含む購入データを分析し、本当にビールと一緒に買われるケースが多いかどうかを確かめるのである。結果として、たまごとビールはあまり一緒に買われないうことになれば、次に「たまごと牛乳ではどうか」といって、再度分析するのである。

しかし、データマイニングでは、そんな面倒な手順は踏まない。そのスーパーマーケットにあるレシート単位のデータ（トランザクションデータという）をすべていっぺんに分析にかけてしまい、

あらゆる場合の通りの中から「たまごと牛乳は一緒に買われる率が高い」ということを発見してしまうのである。

これは、データマイニングという分析手法群の中のひとつの例に過ぎず、それ以外にもさまざまな分析手法が存在するのである。

本書はデータマイニングの解説書ではないので、あまり突っ込んだ説明は避けるが、データマイニングは従来の統計解析に比べれば、圧倒的に膨大なデータを対象として、多面的な分析が可能であることから、ビジネスの現場においても、実に多様な活用のされ方をしているのである。

例を挙げると、ダイレクトメールの対象者の絞り込みであったり、推奨商品の選定であったり、クレジットカードなどの顧客離れ（離反）防止のための傾向把握だったり、販売管理や顧客管理など、多方面での活用余地がある。

◆ Web マーケティングのどこがデータマイニングか

実は、これまでに見てきた Web 上での各種の新技术、新サービスの背景には、このデータマイニングの技術が応用されているのである。

レコメンデーションにしても、行動ターゲティングにしても、プログマイニングにしても、すべてはその根底にデータマイニングの分析手法のいずれかが応用されているのである。

そういう意味では、今日の Web マーケティングの進化の背景にはデータマイニングがあり、データマイニングがなければ、これほどの Web マーケティングの進展はなかったといっても過言ではないのである。

もともと統計解析やデータマイニングについての予備知識があれば、レコメンデーションや行動ターゲティングといった Web マーケティングの手法の裏側にデータマイニング手法があることに気づくだらうが、多くの場合、それには気づかないことが多いに違いない。

そこで本章では、Web マーケティングの革新的なテクノロジーと、その背景で動いているデータマイニングの手法との関係性を明らかにしていくことにする。

2. データマイニング手法と Web マーケティングの関係

◆レコメンデーションを支える技術

Chapter 1 で取り上げた「レコメンデーション」というテクノロジーを支えているのは、データマイニングのアソシエーション分析という手法である。特にレコメンデーションの場合には、マーケットバスケット分析といわれる手法がベース・テクノロジーとして活躍している。アソシエーションとは相関関係のことである。マーケットバスケットとは、その名のとおり、スーパーマーケットなどで、ひとりの消費者が持つ買い物カゴ（バスケット）に、どんな商品と一緒に放り込んでいるか、という分析である。

現在の小売店ではほとんどの店舗で POS システムを導入している。POS においては、レジで精算した顧客ごとの買い物内容の情報がトランザクション・データとして蓄積されている。そのデータをマイニングすることで、「何と何と一緒に買われることが多い」ということがわかる。

この分析を使えば、レコメンデーションに応用できる。たとえば、「ビールとおむつは一緒に買われることが多い」ということが、マーケットバスケット分析の実際の事例としてよく取り上げられる。実際にそうなのだ。奥さんは、荷物としてかさばるおむつを買うことを敬遠し、お父さんに「仕事帰りにおむつを買ってきて」とお願いする。お父さんはせっせとスーパーマーケットにおむつを買いに行き、ついでに自分のためのビールを買うという消費行動をとる（らしい）。

この「おむつとビールは一緒に買われる」ということが明らかならば、「おむつ」を買いたいとサイトにアクセスしたユーザーに対して、「ビールも一緒にどうですか」とレコメンドすれば、ついで買いを誘うことができるという寸法だ（実際問題として、おむつとビールが一緒に買われるのは、荷物になることを嫌がる奥さんの思惑なので、荷物を運ぶ必要のないWebショップには通用しないが）。

こうした分析を膨大なデータに基づいて実施することで、「何かを購入したユーザーには、××をレコメンドすべし」というルールを発見することができるのである。

◆行動ターゲティングを支える技術

行動ターゲティングの基本は、ユーザーがどんな閲覧履歴なり、検索履歴を持っているのかを前提として、「こういう閲覧履歴や検索履歴を持っているユーザーは、××に興味・関心をもっている」ということを判断し、その「××」についての広告をプッシュする、というものである。

具体的な例でいえば、あるユーザーが、過去3日間に、自動車に関するサイトを10回閲覧し、かつ検索サイトで「自動車」というキーワードで6回も検索をしている、ということがわかれば、「そのユーザーは自動車に関心が高い」と判断して、そのユーザーがADネットワーク内で、自動車と無関係のサイトを閲覧しているときにも「自動車の広告を配信する」という仕組みである。

この行動ターゲティングを実現するためのデータマイニングの手法が最適化アルゴリズムといわれるものである。

最適化アルゴリズムとは、ある目的に対して、もっとも合致するパラメータを効率的に導き出す手法である。たとえば、「過去1週間の間に10回、自動車サイトを見た」かつ「過去1週間以内に5回不動産サイトを見た」ユーザーに対して、自動車の広告と不動産の広告のどちらを優先して表示させるか、という判断は一概には決められない。このような場合において、最適化アルゴリズムは役に立つ。

まず、どのジャンルの広告を出稿するかについて、「クリック率を最大化する」などのように目的を決める。さらにパラメータの値によって、自動車の広告にもなり得るし、不動産の広告にもなりえるモデルを設計しておく。そうしておいて、最適化アルゴリズムを適用すると、リアルタイムにクリック率を観測しながら、クリック率を最大化すると思われるパラメータの値を自動的に算出してくれるのである。

◆ログマイニングを支える技術

ログマイニングは、Webクローリング技術とテキストマイニ

ングの技術が融合することで実現した、新しい Web マーケティングのテクノロジーである。

Web クローリング技術とは、ロボット型の検索エンジンであるクローラー（ロボット、スパイダー）と呼ばれるプログラムが、Web サイトやブログなどネット上に公開されているページのリンクをたどりながら情報を収集する行為のことである。

その Web クローリングによって収集された情報は、データベース化され、そのデータベースをテキストマイニングという方法によって解析することにより、どこの Web ページにどんな趣旨・傾向の情報があるかを把握することができるのである。

ここでいうテキストマイニングが、まさにデータマイニング手法のひとつである。

本来、データマイニングが取り扱うデータとは数値データなのであるが、テキストマイニングにおいては、いったんテキストデータを数値化し、その後に数値を分析することが基本である。

ただし、日本語は英語などに比べると少々特殊性があるために、そうそう簡単ではない。

たとえば、英語の場合には単語単位でスペースに区切られているため、分解して分析することが比較的簡単である。しかし、日本語は単語と単語を「てにをは」などの助詞でつなぎ、ひとつの文章として構成されているため、文章を「意味を持つ最小の単位に分割する」ことからはじめなければならないのである。これを専門用語では「形態素解析」という。これが非常に難しいという側面がある。

また、適切なテキストマイニングを行うためには、単語ひとつ1

つの意味を明確に分類しておく必要もあり、前提として活用する「辞書」の充実度も重要なポイントとなる点にも注意が必要である。

◆ LPO を支える技術

LPO (Landing Page Optimization) は、端的に言えば、Web ページのコンテンツを最適化することを意味する。

さまざまな検索エンジンがあり、ユーザーのどんなキーワードの検索によって、結果として自社サイトのどのページが最初に閲覧されるかわからない。たとえば、A というキーワード検索によって、自社の B1 というページにランディングされると、A に関する情報が記載されていないし、A について記載されたページへのリンクも用意されていないために、すぐに離脱してしまい、自社に有効なコンバージョンに至らないということもあり得るのである。

自社の目的どおりに A という検索キーワードから、A1 のページにランディングしたとしても、そのページのレイアウトやテキスト情報 (キャッチコピーなど) が不適切 (そのユーザーに響かない) なために、結局離脱されてしまうということも多々ある。

LPO は、そうした機会損失を防ぐための Web マーケティングのテクノロジーなのである。

たとえば、ユーザーの属性や特性に応じて、複数の“見せ方”が用意されているような場合に、どの“見せ方”がより効果を上げるかについて、最適な組み合わせを導き出すことができるのである。

この LPO を可能にするデータマイニングの手法が「実験計画法」といわれるものである。

実験計画法とは、もともと農業試験用に開発された分析手法なのであるが、要はある目的に向かって、複数のプロセスがあり、かつそれぞれのプロセスにおいて複数の選択肢があるような場合に、どの組み合わせが最適なのかを明らかにするために活用される分析手法である。

ある目的の達成に対して影響を及ぼしていると思われる要因とは何か、またその要因による影響・効果がどの程度なのかを、できるだけ少ない実験回数で分析するための手法なのである。

たとえば、あるサイトのあるページのコンテンツとレイアウト構成に関して、画像 (A) が 2 通り (A1、A2)、キャッチフレーズ (B) が 2 通り (B1、B2)、配置 (C) が 2 通り (C1、C2) 候補があったとする。

このとき、 $2 \times 2 \times 2 = 8$ 通りの組み合わせがあるわけだが、この中から、もっとも効果の高い最適な組み合わせを見つけることを考えるのが実験計画法である。

これらすべての組み合わせを列挙すると表①のようになる。そして、実験計画法の考え方を利用することによって、表②のように、4 通りの組み合わせだけで計測を行えばすむということである。

表②を A について見ると、A1・A2 のどちらの画像についても、B・C 条件 (水準) として B1 と B2 を 1 回ずつ、C1 と C2 を 1 回ずつ設定している。そのため、A1 と A2 の画像について、それぞれ 2 回ずつ計測が行われていて、その平均値や分散を比較することはフェアであると考えられる。表②を B・C に着目してみても、同様のことがいえる。

このように、いくつかの要因について、それぞれにいくつかの

水準が考えられるとき、そのすべての組み合わせについて測定を行うのではなく、各要因の水準を比較できるのに最低限の測定回数で抑えることが実験計画法により可能となるのである。

LPOの実現においては、この実験計画法の概念が大きく貢献している。

■実験計画法・例

表①

No.	A	B	C
1	1	1	1
2	1	1	2
3	1	2	1
4	1	2	2
5	2	1	1
6	2	1	2
7	2	2	1
8	2	2	2

表②

No.	A	B	C
1	1	1	1
2	1	2	2
3	2	1	2
4	2	2	1

3. データマイニング活用の可能性

◆応用範囲は Web だけじゃない

そもそもデータマイニングという分析手法群は、Web のために開発されたものではなく、リアルなビジネスにおけるマーケティング・データを供する目的で開発されたものである。

その手法がたまたま Web マーケティングで応用活用され、さまざまな Web の革新的テクノロジーのベースとなっているのである。

実はデータマイニング自体の活用事例はリアルビジネスのほうが圧倒的に多い。

本章で最後に取り上げている株式会社データフォーシーズは、企業のさまざまなマーケティング・ソリューションのためにデータマイニングの分析手法を提供している企業である。

もしデータマイニングに興味をもったのなら、リアルなビジネスにおいて、どんな活用が可能なのか、ということについても、いろいろと事例を探してみることをお勧めする。

きっと、Web 以外のビジネス・シーンにおいても、役立つ使い方を見出すことができるはずである。

02

データマイニングで
飛躍する企業

データマイニングで飛躍する企業

Company
1

株式会社
データフォーシーズ



「データマイニング」を日本に浸透させる 強い使命感と志を持ったベンチャー企業

I. データマイニングに特化した事業展開

データフォーシーズが自らに課す使命

MBAに象徴される科学的アプローチによる企業経営の推進は、日本を含む世界中の企業に瞬く間に広がった。だが、企業内外の関係やしがらみに支配された経営を行ってきた日本企業は、欧米に比べ立ち遅れていると言わざるをえない。

また、企業体質だけではなく、今日の企業を取り巻く経営環境には多くの情報（データ）が存在し、従来の数字情報からテキスト・ビジュアル情報に至るまで、そのデータ量・複雑さは増すばかりである。データを企業経営に活かすことは、技術的にも容易なことではなくなってきている。

こうした環境に対応すべく、株式会社データフォーシーズは、企業のもつさまざまなデータを解析することで、企業側が必要とする情報を抽出することができる「データマイニング」に特化したビジネスを展開している。膨大な情報量と対峙するための専門スキルを備え、またマーケティング・ファイナンスといったビジネスにも精通し、氾濫する情報から企業を成功に導く文字どおりの宝となる情報を探し出し、ビジネスをより一層加速させること

を使命に掲げ、勘や経験に頼らない意思決定をサポートしている
のである。

現在、「マーケティング」「CRM」「金融系リスクマネジメント」
分野を中心にデータ解析業務をアウトソーシングで請け負いな
がら、その適用領域を拡大させている。

Ⅱ. 既存データとネットデータを融合させたマイニング Web ログ解析にデータマイニング技術を応用する

Web サイトは今や企業にとって非常に重要な顧客との接点と
なっていると言っているだろう。しかし、多くの企業では開設す
ることをゴールと設定し、その後はただ情報を更新するだけとな
っているのが現状だ。Web ログを分析するサービスは一般的になり
つつあるが、分析にデータマイニング技術を使うことでこんな発
見があるということを示す、ひとつの事例を紹介する。

ある高額商品メーカーでは、会員制の Web サイトを運営してお
り、その会員数は 50 万人超、月間 PV は 100 万超という膨大なも
のになっていた。それだけ多くの顧客、または顧客予備軍と接触
する媒体となった Web サイトも、事前に策定したスケジュール
に則ってコンテンツを作成するという非流動的な動きしかとれず、
効果検証も一切なされていなかった。

こうした状況を変えるため、データマイニング技術を用いた分析
および効果検証を行うことになったのである。顧客との関係性の
維持・構築を目的としていたサイトだが、閲覧者が商品を購入し
て初めて利益を生むことに変わりはない。膨大な訪問者の中、い
かにして潜在的な購買層・優良顧客予備軍を探し出し、Web サイ

トから販売店に誘導して売り上げを拡大するかということが課題のひとつとなっていた。そこで、データマイニングの手法を駆使し、サイト行動のパターンをクラスター分類し、どんな行動をとる顧客が購買しやすいのか、またその後優良顧客となるポテンシャルを持ち得るのかを分析して、顧客判別モデルを作り上げたのである。

さらに、データフォーシーズではアンケート調査分析も得意分野として、それぞれの分類された顧客群（顧客クラスター）がどのような生活行動・消費行動をとっているかを知るためにインサイト調査を実施した。結果、分類された顧客群ごとに広告のテストを変えるとといった個別のアプローチができるようになり、“one to one”に近いプロモーション展開ができるようになったのである。

ある特性を持つ顧客のグループは、入会後の継続閲覧率が低いということがわかり、その顧客群を想定したコンテンツを新たに揃えるようにするなどして、分析結果に基づいたサイト作りを社内で徹底するようにした。クリエイティブとマーケティングチームが共通のゴールに向けて動くことができ、両者の関係性がより密接になったという。

Ⅲ. 顧客の CRM 戦略をデータマイニングで支援する 現場が使えるシステムを構築する

成熟した市場環境下において、従来の新規顧客開拓よりも既存顧客との関係強化を経営戦略上重視する企業が増加している。CRMは顧客当たりの生涯価値を最大化するための戦略であり、個別の施策としては、顧客ポートフォリオの設計や管理を基盤として、

解約防止・クロスセル、アップセル・プロモーション、価格政策等がある。

以下では、サービス業のCRM戦略の支援を目的としたデータマイニングの事例を紹介しよう。

全国の支店でサービス業を展開しているA社では、経営戦略として顧客の解約を防止することを重要視している。販売管理に関する情報は、データベースに蓄積されており、本プロジェクトではこの膨大なデータから解約しやすい顧客を予測し、解約防止活動を支援することを目的とした。

手始めに、社内に分散している顧客情報を統合したデータマートの作成を行った。顧客ごとの基本属性・さまざまな期間の売り上げや契約に関する履歴・支店属性・サービス担当者属性などの情報を統合/整理した情報系のデータベースである。

次に、作成したデータマートを使い、将来の解約を予測するモデルの構築を行った。実際に解約した顧客の特徴的な属性や行動を整理・統合化することにより、解約しやすい顧客像をモデル化するのである。解約予測モデルを構築することで、解約のシグナルをいち早く察知することができるのだ。

最後に、作成した解約予測モデルを販売管理システムに組み込み、各支店で現場の担当者が解約予測情報にアクセスできる仕組みを構築している。各担当者は、この情報を使い、解約確率が高いと予測された顧客に、重点的なアフターケアや価格やプロモーション施策の見直しなどのアクションを効率的に推進することができるようになるのだ。

データフォーシーズでは、上述した解約防止以外にも、さまざま

まな領域でデータマイニングの手法を駆使したサービスを提供し、企業のCRM施策の現場で大きな力を発揮している。

IV. ファイナンス分野における データマイニング技術の活用

数理科学の力で金融機関が抱えているリスクを計量する

昨今のサブプライム問題を例に挙げるまでもなく、国際化・自由化が進展する中で金融機関が抱えるリスクは深刻さを増すばかりである。それにともない、リスク管理の重要性は高まる一方であるが、深刻さと比例するかのようにその内容は複雑化してきている。

データフォーシーズでは、データマイニングで培ったノウハウを武器にリスク計量に挑んでいる。現在、金融機関の住宅ローン業、リース業といった個別具体的なリスク計量に携わっているが、ここでは銀行本体が抱えるオペレーショナルリスクの計量について紹介しよう。

オペレーショナルリスクとは、社員の不祥事、業務プロセスやシステムの不具合、犯罪、天災等により生ずるリスクである。このようなリスクは減多に起きることはないが、発生すると甚大な損害額になるという特徴を持つ。また、過去の事例が少ないこともリスク計量を困難にさせる一因となる。

リスク管理とは文字どおり「リスク」を「管理する」ことなのだが、オペレーショナルリスクに関しては、以下の3つのプロセスを実施する。まずどのようなリスクが存在するのかというリスクの想定を行い、そのリスクが発生した際の損失額を見積もり、最後に