

マーケティング分野における分析システム 日本ユニシスの取り組み

The Analytical System in Marketing Field
The Measures of Nihon Unisys

松 田 芳 雄

要 約 日本ユニシスでは、マーケティングの分野を中心に分析系のシステム開発に取り組み各種ソリューションを提供してきた。データベースマーケティングを実現する IMPACT DM/MA, 数値データと文字データのデータマイニングを行う MiningPro 21, 顧客の声の管理と分析を行う CVPro などである。それらは単にソフトウェアの機能を提供しているのではない。事前のテスト分析による提案, システムの導入, 分析のサービス, 導入後のコンサルティングなどの提供により効果を上げている。本稿では、日本ユニシスが提供してきた分析系ソリューションの開発目的とその考え方を紹介する。

Abstract Nihon Unisys tackles the system development of an analytical system focusing on the marketing field, and provides various solutions, including IMPACT DM/MA to realize database marketing, MiningPro 21 to perform (numerical) data mining and text mining, and CVPro to manage and analyze the customer's voice. The above software do not only provide the software functions, but also achieve the effect by providing the proposal based on pre test analysis, introduction of the system, analytic service, post implementation consulting. This paper introduces the purpose for and the concept under which the analytical solution that Nihon Unisys has offered was developed.

1. はじめに

日本ユニシスでは、15年以上前からマーケティング分野の分析システムに取り組んでいる。最初はデータベースマーケティング分野での分析システムの開発であった。日本のデータベースマーケティング業界では1980年代後半に業務処理のシステムの整備が終り、購買データを中心とする顧客情報の蓄積が始まった。これを契機にデータベースマーケティングにおける分析系の研究が始まり、1992年にデータベースマーケティング・システム IMPACT DM/MA が開発された。データベースマーケティングの分析においてはデータマイニングの機能も必要であり、個別に開発してきたツールを集大成して、1999年にデータマイニング・ソリューション MiningPro 21 の提供を開始した。データマイニングでは、顧客が買うか買わないかの購買可能性は予測できるが、なぜ買うかなぜ買わないかの購買の理由は分からない。これを解析するために、顧客の声に対する文書マイニングの研究が始まり、2003年には顧客の声システム CVPro を発表した。

本稿では、IMPACT DM/MA, MiningPro 21, CVPro がどのような考えのもとに開発されたかを紹介し、日本ユニシスのマーケティング分野における分析システムに対する考え方を述べる。

2. データベースマーケティングへの取り組み

企業に蓄積された顧客情報を利用して行うデータベースマーケティングには二つの目的があ

る。一つは、将来のマーケティング戦略を策定するための戦略面での利用である。もう一つは、個々のプロモーションの効率向上を狙う販促面での利用である。

2.1 戦略面での顧客情報の利用

データベースマーケティングを最も先進的に実践している業界は、店舗を持たずに保有する顧客リストをもとに商品などの販売を行うダイレクトマーケティング業界である。この業界を例に顧客情報の利用について考える。ダイレクトマーケティングでは、住所も名前も分からない潜在顧客に対し、新聞や雑誌の広告、新聞折込チラシなどで新規顧客を獲得して顧客リストを作成する。この顧客リストをもとにカタログやDM（ダイレクトメール）を送って2回目、3回目の購買を促し継続購買を得ることによって収益を上げている。

1) データベースマーケティングの顧客構造

ここでダイレクトマーケティング業界の顧客の構造を考える。まず新規顧客獲得の費用を見る。一般に、新規に獲得した顧客からは利益が上がっていない。たとえば、新聞折込チラシで新規顧客獲得を行った場合の収益を考える。仮に1部10円のチラシを100万部新聞に折り込んだとすると総費用は1,000万円である。このときのレスポンス率が0.15%だとすると1,500人の顧客を獲得することができる。このチラシでは商品の販売も行っており、1人当たり平均6,000円の売上があるとすると総売上高は900万円となる。商品原価率が50%なら粗利益は450万円である。チラシ折込費用が1,000万円であるので利益はマイナス550万円となる。すなわち、1,500人の新規顧客を獲得するために550万円の費用を使ったことになり、新規顧客1人当たり獲得費用は3,667円（ $= 550 \text{万円} \div 1,500 \text{人}$ ）である。顧客を1人新規に獲得するたびに3,667円ずつの赤字であり、売れば売るほど赤字が膨らんでいることになる。ただし、一度獲得した顧客に対してはDMなどで次の購買を促す販促が行われる。1人当たり3,667円の費用を掛けて獲得した顧客1,500人に1部当りの費用が100円のDMを発送したとする。このDMの総費用は15万円である。レスポンス率が30%で購買顧客1人当たり平均1万円の売上があるとすると、総売上は450万円（ $= 1,500 \text{人} \times 30\% \times 1 \text{万円}$ ）で粗利益は225万円（商品原価率50%）となる。総利益は210万円（ $= 225 \text{万円} - 15 \text{万円}$ ）で1人当たり利益は1,400円（ $= 210 \text{万円} \div 1,500 \text{人}$ ）である。この1,500人の顧客には新規獲得時に1人当たり3,667円を掛けているので1,400円の利益があったとしても累積利益は2,267円（ $= 1,400 \text{円} - 3,667 \text{円}$ ）でまだ損失が出ている。さらに同様のDMをもう一度送って同じ利益を得たとしても867円（ $= 1,400 \text{円} + 1,400 \text{円} - 3,667 \text{円}$ ）の損失である。3回目のDMでようやく533円（ $= 1,400 \text{円} + 1,400 \text{円} + 1,400 \text{円} - 3,667 \text{円}$ ）の利益が出る。以上のようにダイレクトマーケティングのビジネスでは、リスト上の既存顧客からの繰り返し購買によって収益が確保されていることがわかる。

2) 新規顧客獲得と既存顧客維持のマーケティング

このようにデータベースマーケティングでは新規顧客獲得と既存顧客維持の二つのマーケティングが行われている。一般的に新規顧客からの収益に比べて既存顧客からの収益は高い。しかし顧客がいつまでも購買を続けることはなく、いつかは購買を停止して顧客は離脱する。つまり顧客維持率を100%にすることは不可能で既存顧客は減少していく。この減少分を新規顧客の獲得で補わなければ売上が減少し事業は縮小する。ある期間におい

て新規に獲得した顧客と離脱した顧客の数が等しければ現状維持である。離脱顧客よりも新規獲得顧客の数の方が多ければ事業は拡大する。逆に新規獲得顧客よりも離脱顧客の方が多ければ縮小である。データベースマーケティングを実践している企業は、新規顧客の獲得数、既存顧客の数、離脱顧客の数の微妙なバランスで経営が成り立っている。このバランスを最適な状態に保つために、マーケティング費用を新規顧客獲得と既存顧客維持の二つのマーケティングに適切に配分することが、顧客情報の戦略面での利用の目的である。

3) 顧客のライフサイクル (RFM 分析)

マーケティング費用の適切な配分のために顧客のライフサイクルを考える。企業との関係において顧客にはライフサイクルが存在する。購買のない見込顧客が1回購買すれば新規顧客となる。さらに購買を繰り返せばリピート顧客となり、購買が定常化すれば優良顧客である。顧客が永久に購買を続けることはなく、いつかは購買がなくなり離脱顧客となる。離脱の過程で一時的に購買がなくなったと思われる顧客を休眠顧客と呼ぶこともある。顧客の休眠や離脱は新規顧客、リピート顧客、優良顧客のどの段階からも起こりうる。これらの顧客の状態推移を顧客のライフサイクルとよぶ。顧客がライフサイクルのどの状態にあるかを把握するためにRFMという指標が用いられる。R (Recency) は最新購買日で顧客との取引の最も新しい日付である。F (Frequency) は累積購買回数で顧客とのこれまでの取引の回数、M (Monetary) は累積購買金額で取引の累積金額である。このRFMを使って顧客のライフサイクルを推定することができる。たとえば、RとFの組み合わせでマトリックス (R×F 管理表) を作成する。Rが1年以内で購買回数1回の顧客は今年獲得した新規顧客である。Fが2回以上ならば複数回購買したリピート顧客であり、さらにFが多ければ優良顧客である。また、Rが1年を超えている顧客は今年の購買がなく、休眠しているか離脱した顧客である。Mは1回の高額購買で優良顧客になることがあるため、顧客を評価する場合には一般的にRとFが使われているが、単価の安い商品を取り扱っている場合は購買回数が多くても売上に寄与していない場合があるためRとMが使われることもある。

R×F 管理表によって新規顧客、リピート顧客、優良顧客、離脱顧客 (休眠顧客) のそれぞれにどれだけの顧客が存在しているかを把握することができる。さらに新規顧客からリピート顧客や優良顧客への推移の状況、優良顧客の維持の状況から既存顧客の減少数を分析することができる。これから必要な新規顧客の獲得数も判明する。既存顧客の維持と新規顧客の獲得の二つのマーケティングに費用をどのように配分するかの戦略を立案することができる。データベースマーケティングはこのライフサイクルで顧客を管理し戦略の立案を行う手法である。

4) データベースマーケティングによる顧客戦略 (LTV 分析)

売上に目標を設定している企業は多いが、顧客の数に目標を設定している企業は多くはない。企業の売上の源泉は顧客であり、R×F 管理表のどのセルにどれだけの顧客がいるかによって売上が決定される。データベースマーケティングでは、売上目標が示されたときそれを顧客戦略に展開して顧客の数と構成を計画する。すなわちR×F 管理表の各セルの必要な顧客数が計算され目標が設定される。たとえば、売上目標達成のためには購買のある稼働顧客が何人必要か、離脱する顧客を補填するために新規顧客を何人獲得するか、または事業拡大のために新規顧客を何人獲得するか、あるいは将来の優良顧客を10%増

加させるためには各顧客層の維持率をどれくらいするかなどの目標の設定が行われる。さらにその顧客戦略を実現するために個別の販促計画が立てられる。

顧客目標の設定にはLTV (Life Time Value ; 顧客生涯価値) 分析が使われる。顧客の新規獲得を全くしなかった場合、現在保有の既存顧客から将来に渡ってどれくらいの売上または利益を確保できるかで顧客リストを評価する方法である。たとえば現在の顧客リストが1年後にはどのような顧客数と構成になっているかを計算することができる。RとFのマトリックスで顧客が管理されているとすると、R1年以内のF1回顧客は1年後にはF2回、3回、...、と購買回数を増やすか、離脱してR1年超F1回顧客に推移する。R1年以内F2回顧客も同様に、R1年以内のF3回、F4回、...、またはR1年超F2回顧客に推移する。他のR1年以内顧客も同様である。1年間購買のなかったR1年超F1回顧客は、そのままR1年超F1回顧客にとどまるか、購買が復活してR1年以内F2回、3回、...の顧客に推移する。各顧客のセル間の推移確率は過去の顧客の購買データから推定することができる。このようにして1年後の顧客の状態をシミュレーションにより計算する。これに顧客の平均購買単価を掛ければ1年後の顧客リストから期待できる売上が求められる。この売上が利益に換算したものがLTVであり、現在の顧客リストの評価値となる。シミュレーションにより求められた1年後の顧客の状態により新規顧客の獲得数が計画される。また、リピータや優良顧客の減少幅の大きいところがあればその部分に対する販促が強化される。

5) 顧客マネジメント中心のマーケティング

一般に売上の減少よりも顧客の数の減少の方が早く現れる。購買の多い優良顧客よりも購買の少ない下位顧客の方から先に離脱が始まるため、売上の減少に気が付いたときには顧客はそれ以上に減少していることが多い。図1はある企業の顧客数の推移と売上高の推移の関係をシミュレーションした結果である。すなわち、顧客の新規獲得を行わなかった場合における、既存顧客の顧客数 (R1年以内の稼働顧客数) と売上高の時系列推移を試算したものである。顧客数は1年後には79.2%、2年後には55.7%、3年後には46.2%、

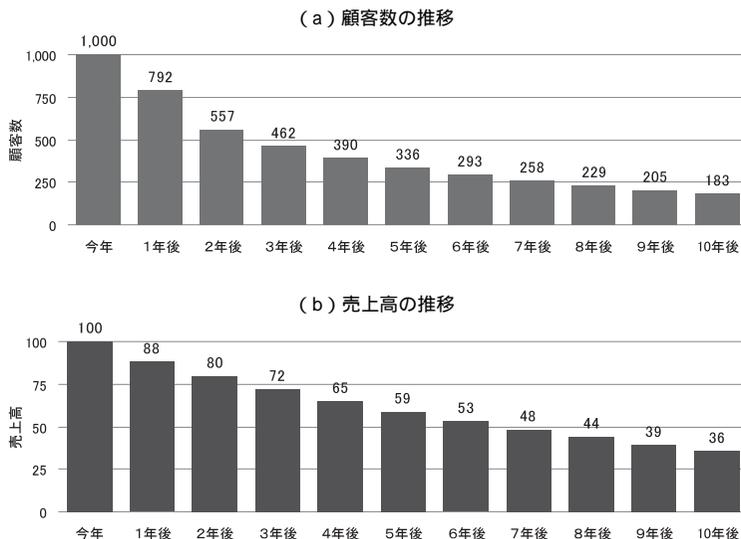


図1 顧客数の推移と売上高の関係

...と推移している。これに対し売上高は1年後88%、2年後80%、3年後72%...と推移している。2年後に売上高が2割減少したときには、顧客の数は44.3%も減少しており、減少率は売上高の2倍以上になっていることがわかる。半分近くに減った顧客を元に戻すことは容易なことではなく、新規に事業を起こすことと同じくらいの努力を必要とする。このように売上高だけ見ているとこの顧客の減少に気が付かないことがある。顧客がいなければ売上は上がらない。顧客一人ひとりの購買単価や購買頻度を上げて顧客維持率を100%にすることは不可能である。いつかは売上ゼロの日が来ることになる。顧客数は売上高の先行指標になっており、各種販促の効果は顧客の数に反映され、施策の失敗は顧客の数に表れる。

ある企業は売上が低迷してきたときに、新規顧客獲得のプロモーションの見直しを行った。少しでも採算を良くするために単価の高い商品で新規顧客獲得を行った。その結果新規獲得のプロモーション自身の採算は改善されたが、獲得した顧客からの2度目以降の継続購買がなく、2年後の業績が低迷してしまった。原因が判明しないままさらに新規獲得に重点投資したが結果、業績は悪くなるばかりであった。ダイレクトマーケティングが顧客の継続購買を前提として成り立っている限り、単発の販促プロモーションの結果ではなく、2年後、3年後の顧客の状態を予測した顧客マネジメント中心のマーケティングが必要である。

6) リスト診断

データベースマーケティングでは、企業の売上目標は顧客数の目標に展開され各種販促計画が策定される。また、月次や四半期ごとに目標と現在の顧客数との乖離がチェックされ問題があれば販促計画が見直される。前述のように顧客数は売上高の先行指標になっており、問題のあるプロモーションがあればその部分の顧客数が減少する。個々のプロモーションの評価ではなくプロモーション全体が順調に推移しているかを監視しなければならない。施策の失敗はそのとき直ぐにはなく1年後や2年後、3年後に大きく影響することが多い。現在の顧客リストが将来どれくらいの価値を持つかを計算して将来計画を立てる必要がある。図1から現在保有する顧客リストからの1年後の売上高は88%に減少することが分かる。この値を顧客リストの現在の評価値と考えることができる。すなわちこの顧客リストは100点満点で現在88点である。顧客数の増減と顧客リストの評価値を毎日監視して、各種のプロモーションや施策が間違いなく進行していることを確認することが重要である。リスト診断はこのように毎日の顧客の数とその変化を監視し、顧客リストの評価を行うシステムである。

図2はリスト診断システムの画面例である。毎日、昨日の業務終了後の顧客の数を集計して増減を表示するようになっている。万一顧客数が減少しているときには日本語で警告を出し、見落としのないようになっている。また、前述の100点満点のリスト評価値を表示して、顧客リストの健全性を確認できるようになっている。この画面は、毎朝社内イントラネットのWeb画面に自動的に表示されるようになっている。経営者から担当者まで大勢の目で顧客リストを監視し、共通の認識のもとでそれぞれの業務を遂行することができる。一人の担当者による定期的な顧客リストの分析はおざなりになることがあるが、半ば強制的に顧客リストの状態を把握できるこのようなシステムがデータベースマーケティングには必要である。



図2 リスト診断システムの画面例

2.2 販促面での顧客情報の利用

ダイレクトマーケティング業界では、中堅企業で100万件前後、大手企業で約1,000万件的顧客リストを保持している。その内の30%前後の顧客を選んでカタログやDMを送付して商品を販売している。たとえば100万件的顧客リストの中には購買可能性の高い顧客もあれば全く見込みの無い顧客も混在している。カタログやDMを送付するためには保有の顧客リストから購買可能性の高い顧客を識別する必要がある。

1) RFMによる顧客評価

顧客の購買可能性を推定する最も簡単な方法は最新購買日Rによる方法である。Rが新しい顧客はその企業や商品に対する記憶も新しく印象も強いことから再購買の可能性は高い。Rの次には累積購買回数Fが有効な指標である。累積購買回数Fが多い顧客はその企業との取引も長く、商品への理解もあり、購買も習慣化している場合が多く再購買の可能性は高くなる。通常はRと組み合わせで利用される。すなわち、RとFのマトリックスを作成し、各セルの過去のレスポンス率や平均金額などの実績を調べ、購買可能性の高い顧客を抽出する。特定の商品を販売したい場合は、顧客が過去に購買した商品の情報が使われることもある。過去に顧客が購買した商品分野ごとにRとFのマトリックス表を作成しその商品の購買が多いセグメントを抽出している。さらに、性別や年齢などのデモグラフィックスの属性、新規獲得の時期、どのプロモーションで新規獲得されたか、過去にどのようなカタログやDMが送られ販促が行われていたか、春夏秋冬のどのシーズンに購買が多いかなどさまざまな要因で顧客の評価が行われている。使用される要因が多ければ顧客の購買可能性推定の精度は向上するが、集計ベースの分析では組み合わせの数が増えて人手で処理することは困難になる。そこでデータマイニングを利用したスコアリングという方法が利用される。

2) スコアリングによる顧客評価

スコアリングは、さまざまな要因を数値化し、顧客に購買可能性を表す得点を与える方法である。通常は期待確率や期待金額などわかりやすい数値で表現される。期待確率は顧客の購買可能性を確率で表現したものである。たとえば期待確率 5% の顧客が 100 人いたとするとその内の 5 人から購買が見込めるということである。期待金額は顧客の購買可能性を金額で表したものである。たとえば期待金額 1,000 円の顧客が 100 人いれば購買する顧客もしない顧客も全部まとめて平均金額が 1,000 円になる顧客ということである。顧客を確率や金額で表現することによりいくつかの利点がある。一つは、受注の予測を簡単に行うことができるようになる。DM を送付するために抽出した顧客が 10 万人あるとすると、この顧客の期待確率の平均値を求めることで受注人数が予測できるし、期待金額を合計することにより受注金額を予測することが可能になる。また期待確率や期待金額の値の大きい順に顧客を選択して、DM などで販促することによって容易に受注の最大化を図ることができる。過去の事例から、休眠顧客以外の一般顧客では従来の RFM の方法に比べてスコアリングによる方法は 3% から 7% 程度の受注増大が期待できる。休眠顧客では 3 倍から多ければ 5 倍程度の効果が出ている。スコアリングは顧客選択効率の高い手法である。販促の費用が 1 人当たり 800 円掛かるとすると、期待金額 800 円以下の顧客からは利益は出ない。採算範囲内にある顧客がどれだけいるかを見極め、最適 DM 部数の設定や販促戦略を立てるためにもスコアリングは有効な手法である。

2.3 日本ユニシスのデータベースマーケティング・システム IMPACT DM/MA

日本ユニシスでは、1980 年代の後半からデータベースマーケティングのためのシステム開発を行ってきた。1980 年代後半は日本のダイレクトマーケティング業界の基幹系システムの整備が終り、購買データをはじめとする顧客情報が蓄積され始めた頃である。当初は、少しでも購買可能性の高い顧客を抽出して DM 送付の効率化を図る販促面での利用から始められた。1992 年には日本で初めてスコアリングによる顧客選択システム IMPACT DM/MA を発表した。顧客に購買可能性を表す得点を与え、売上の最大化と受注の予測を目的としたシステムである。1994 年には IMPACT DM/MA (DSS) の開発により、販促面だけではなく戦略面でも顧客情報を活用することを可能にした。その後データマイニング機能 IMPACT DM/MA (PLUS) を充実させ、業界で唯一のデータベースマーケティングのための統合分析システムとなっている。また、2005 年 4 月にはリスト診断機能や受注予測機能、売上速報機能などを付加し、従来機能も一新した MA 3 を発表して常に新しい技術を提供し続けている。

3. データマイニングへの取り組み

データマイニングは、大量データの中から規則性や関連性など意味のあるパターンを自動的に抽出する手法である。多変量解析などの統計的手法から人工知能の機械学習分野の手法までさまざまな手法が用いられている。1990 年代の中頃から注目され始め、一時はブームにもなったが導入に失敗した企業も多い。昨年 (2004 年) 4 月より 9 月まで、日本ダイレクトマーケティング学会において通信販売企業 8 社が参加して、顧客情報活用のための研究プロジェクトが結成された。そこで何故データマイニングが使われていないかが議論されている。一方、うまく活用して効果を上げている企業も多い。データマイニングの有効な利用分野と導入に失敗

した理由などを考察する。

3.1 データマイニングの利用分野

昨年の日本ダイレクトマーケティング学会の研究成果に基づいて、データマイニングが効果的に使われている利用分野を検討する。昨年の研究参加企業は全て通信販売企業であり、データマイニングは主にカタログやDMを送付する顧客を選択する業務に利用されている。ここでもその分野に限定してデータマイニングがどう使われているかをみる。

1) 休眠顧客の掘り起こしに利用

A社は、Rが1年以内の稼働顧客には顧客との関係維持のため全員にカタログを送付している。Rが1年超の休眠顧客はRやFで選択を行っているが、データマイニングを利用したスコアリングによる方法を試行して2倍近い受注率を得ている。これは売上の確保というよりも稼働顧客を増やすことが目的である。

B社も、RFMの上位顧客には無条件にカタログを送付しているが、下位顧客に対してレスポンス効率改善のためにデータマイニングを利用したスコアリングによる方法を導入し、2割以上の効果を上げている。部数を絞っても前年売上を維持できるくらいの精度を実現している。

C社は、1年前より本格的にデータマイニングのテストを行っているが、ここでも目的は休眠顧客の掘り起こしである。顧客をRFMのセル単位から個人単位に評価することにより効率化を図ろうとしている。データマイニングと従来の方法を組み合わせることにより効率を上げている。

2) 別冊カタログの顧客選択に利用

D社は、カタログ本誌のほかに別冊カタログを20種類近く発行している。以前はRとFのポイントで顧客選択を行っていたが、カタログの数が多く組み合わせも発生するので人手では対応できなくなりデータマイニングを採用している。RFポイントによる方法に比べレスポンスは数パーセント程度改善されているが、レスポンスの改善よりも人手の負荷軽減が目的である。

3) レスポンスの効率化と作業の効率化

以上よりデータマイニングの利用に二つの目的があることが分かる。すなわち、レスポンスの効率化と作業の効率化である。

上にあげたA社、B社、C社はRが1年を超えた休眠顧客に対するカタログ送付顧客の選択にデータマイニングを利用している。いずれも稼働顧客の増加や受注金額の増大を目的としたレスポンスの効率化が狙いである。休眠顧客は最後の購買から時間が経過しておりRFMによる顧客選択の効果が稼働顧客に比べて小さくなっている。特に最も要因効果の高いとされるRが均一化してしまい顧客選択の効果が小さくなっている。他の要因も同様で少ない情報量から顧客の選択を行わなければならない。このような少ない情報量から高効率の顧客選択を行う場合は、各要因の組み合わせや重み付けを厳密に行う必要があり、人の手で行うには限界がある。また、一つひとつの要因の情報量が少ない場合は、少しでも情報量を増やすために要因の数を増やす必要があり、これも人の手で効率化を行うことを困難にしている。そこでデータマイニングのようなデータに基づいて数学的に最適化する方法が必要になる。休眠顧客においては人による方法に比べデータマイニングで2

~3倍の効率を出している例もある。このようにデータマイニングが効果的に利用されている分野の一つは、情報量の少ない中で高い精度の実現を必要とする分野である。

D社は別冊カタログの送付顧客選択にデータマイニングを利用している。従来はRFポイントにより顧客選択を行っていたが、データマイニングの採用による効率改善は数パーセントであり、効率の追求だけが目的ではない。20種類以上ものカタログがあると作業量が増えて人手での顧客選択は困難になる。そこで人が行うのと同程度かそれ以上の精度で作業をシステムにより自動化する必要が生じる。自動化のためには顧客選択基準の標準化や数量化が必要であり、データマイニングでこれを実現している。すなわち、人手の作業をシステムにより自動化するための評価基準の作成にデータマイニングが使われている。

以上より、データマイニングが利用される分野は、情報量の少ない中で効率化が要求される分野と、作業量が多く手作業では不可能な作業を自動化するための基準作りが必要な分野である。

3.2 データマイニングを利用しない理由

前節のようにデータマイニングが効果的に利用されているケースも多いが、失敗したケースも多い。日本ダイレクトマーケティング学会の研究成果から、テスト導入はしたが本格的には導入しなかった、あるいは以前は使っていたが今は使っていない理由を挙げると以下になる。

1) データマイニングに対する誤解と過度な期待

テスト導入時に「目新しい結果が出なかった」「成功例を示せなかった」「分っていた答えしか出てこない」などが、データマイニングを利用していない企業が挙げた理由である。これらにはデータマイニングに対する期待との乖離が現れている。データマイニングを適用すれば、今まで知らなかった新しい事実が発見できるものとの期待がある。そもそも誰も知らない新発見の事実など存在するものなのか、存在したとしてもデータマイニングでそれが簡単に発見できるものなのかは疑問である。データマイニングは1990年代中頃に提唱されたデータ解析ための手法の総称である。その中の個々の手法はそれほど新しいものではない。たとえば、決定木を作成する手法はいくつかあるが、その中でAID(Automatic Interaction Detector)は交互作用を検出する手法として1963年にミシガン大学で開発されたものである。その後CHAID,CART,ID3,C4.5,QUESTなどの手法が開発されたが、それらは適用範囲を質的変数に拡大したり分割の評価基準を変えたりしたAIDの改良版といってもよい。本質的には40年以上前の手法と大きな違いはなく、データマイニングはそれを大量データで行って結果の信頼性を保証しようというものである。決して知らない事実が次々に発見できる魔法の手法ではなく、それを期待して導入すれば裏切られることになる。「データマイニング自身もよく理解していなかった」という意見もあり、データマイニングで何ができてどれくらいの効果が得られるかを正しく理解していないことが利用の促進を阻害する要因の一つとなっている。

2) データマイニングの適用分野と評価

「微妙な効果」「数パーセントしか上乗せ」「レスポンスの差は数パーセント程度」などとデータマイニングによる顧客選択の効果が期待以上ではなかったという理由もある。カ

タログのレスポンスが数パーセントも違えば大きな差があるとも思えるが、RFM 上位の優良顧客の顧客選択は誰がどのような手法で行っても大きな差はない。この部分は情報量も多く顧客の優劣の差が明確で購買可能性の高い顧客を見つけることは容易である。前述のように休眠顧客などの情報量の少ない部分は人手による効率化が難しく、データマイニングの効果が発揮できる領域である。データマイニングの適用領域が適切でないと大きな効果は上げられない。「損益分岐点が低いためデータマイニング導入の経済的効果は小さい」との意見もあり、導入前にはテスト分析等により効果の評価を事前に行うことが必要である。

3) 人材の問題

「強い意志を持った人がいなかった」「使いこなす人材がいなかった」のような人的要因の問題も挙げられている。データマイニングにはモデル化が必要である。たとえば、カタログ配布顧客の選択を行う場合、手法を選択して何を説明変数として何を外的基準とするかなど分析のやり方を分析者が自分で組み立てる必要がある。これを初心者で独学で会得するには時間が掛かり途中で放棄する人も出る。放棄のついでにデータマイニングの評価を“役に立たない”と結論付けてしまう例も多い。データマイニングはソフトウェアやツールを導入すれば直ぐに使いこなせて効果が出るものではない。導入の最初には熟練した指導者が必要である。また、分析者自身の能力の問題もある。問題の本質を理解できるだけの業務知識、データを取り扱うための IT 技術、手法を理解するための少しばかりの数学的な知識などを持ち合わせていなければならない。すなわち、データマイニングの導入には分析能力のある人材とそれを指導する体制が必要である。

4) マイニング用データ作成の問題

「データの作成が手作業で面倒」という、データマイニング用データの作成に関する問題もある。手間だけの問題であればデータ作成の仕組みをシステム化すればよいが、データマイニングの目的によって分析用のデータを作り分ける必要がある。どのようなデータを作ればよいかは分析の初心者には難しい問題であり、人的要因の問題と同様に分析者の能力と指導者に関わってくる。

5) 経営者や上司の無理解、無関心

逆に、データマイニングの導入を推進した企業は「部長の交代」や「部長の関心が高かった」など上司の関心の高さや理解が挙げられる。データマイニングは企業の経営や運営に必ずしも必要なものではないため、経営者や上司の理解が得られないと導入は難しい。データマイニング導入の効果が数値で示され、それを経営者が理解できなければ導入は難しくなる。

データマイニングの一時的な流行に乗り、導入前に効果検証も行わずに、過大な期待を抱いてツールを導入したが、思い描いていた結果が得られず失望し断念しているケースがある。また分析の方法やモデル化の方法が分らず挫折しているケースも見受けられる。これらはデータマイニングやモデリングに対する誤解や無知識によるものである。データマイニング普及のためには、優秀な指導者による正しい教育が必要である。

3.3 日本ユニシスのデータマイニング・ソリューション MiningPro 21

前節で検討したようにデータマイニング導入の阻害要因の一つにデータマイニングを使いこなす人材がいなかったなど人的な問題があった。データマイニングはデータを入力すれば直ぐに効果的な出力が得られるというものではなく、分析の手順を分析者が自分で組み立てるモデル化が必要である。モデル化は分析の初心者が独学で学ぶには難しい面があり、しかるべき指導者が必要である。しかし、学校や各種セミナーなどに通ってこれを取得するのは毎日の仕事を抱えたビジネスマンにはそう簡単にできることではない。初心者にとって分析をナビゲーションしてくれるソフトウェアがあれば理想的である。

日本ユニシスのデータマイニング・ソリューション MiningPro 21 は 1999 年 10 月 19 日に最初のバージョンがリリースされた。一般のデータマイニングツールは手法を中心に構成されているのに対し、MiningPro 21 は問題が中心になっている。マーケティングの世界ではデータマイニングは顧客のランク付けとグループ分けに利用されることが多い。DM を送付する顧客を選択するために見込み度の高い順に顧客を並べ替えたり、同じ購買特性を持つ顧客を一まとめにしてそれぞれのグループごとに販促の企画を立てることなどが行われている。これらを行う担当者はマーケティングの専門家であっても分析業務の専門家でないことが多い。MiningPro 21 はこれらの人達がデータマイニングを簡単に利用できるよう工夫がなされている。ポイントは、①モデル化を容易にする、②操作を簡単にする、③結果の解釈に専門知識を必要としない、の3点である。

① モデル化を容易にする

顧客のランク付けやグループ分けはさまざまな方法があり決まったやり方があるわけではない。分析の目的に従って使用するデータを決め、それを適用する手法を決め、何を出力とするかを定めることをモデル化と呼ぶ。通常は分析者が最も良いと思われる方法を考えてモデル化を行うが、初心者には難しい作業である。MiningPro 21 には過去のデータベースマーケティングで実践された数多くの事例の中から有効と思われる方法がプログラムとして組み込まれている。分析の初心者でも過去の成功事例をそのまま自分の問題に適用できるようになっている。すなわち、MiningPro 21 は問題の解決を中心に考えられたツールとなっている。

② 操作を簡単にする

一般的にデータ分析は、モデルの作成、パラメータの推定、モデルの検証、予測の手順で行われる。通常はこの手順も分析者が組み立てることになるが、MiningPro 21 ではあらかじめプログラム化されており、二つか三つのボタンを押すだけで自動的に実行できるようになっている。分析手順の組み立てとツールの操作に余計な気遣いと労力を使わせないための配慮がなされている。

③ 結果の解釈に専門知識を必要としない

分析の結果を実務に適用して効果を上げるためにはデータから十分に情報を抽出していなければならない。データからの情報の抽出度合いをモデルへの適合度と呼ぶと、データマイニングでは適合度は統計学理論などに基づいた統計量として出力される。この統計量は統計学をある程度理解していないと解釈できないし、それを実務に適用したときにどれほどの効果が得られるかは分からないのが普通である。MiningPro 21 では各種統計量の他に実務に適用したときの効果が分かるような出力を用意している。たとえ

ば、DM 送付顧客の選択を考えると、10% の顧客を選択したら何パーセントのレスポンスと受注金額が得られるか、50% の選択だったらどうなるかなどの図表が出力され、統計学的な専門知識が無くても実用的な見地からモデルの評価が行えるようになっている。

このように MiningPro 21 は実務家を対象としたパッケージである。データマイニング導入の阻害要因に「使いこなす人材がいなかった」という声があった。要員の育成は重要な問題であるが、分析要員の人口を増やすために分析技術の専門知識を持たない一般の利用者にも拡大できるようなソフトウェアからの取り組みも必要である。MiningPro 21 はそのような方針で開発されたソフトウェアである。

4. 顧客の声分析への取り組み

データマイニングを適用することによって個々の顧客の購買可能性を推定することは可能である。その顧客が買うか買わないかを事前に判定することができる。また何を買うかもある程度推定することができる。しかし、なぜ買うのか、なぜ買わないのか、購買の理由までは分からない。ダイレクトマーケティングでは新規に獲得した顧客のうち2回目の購買がある顧客は半分以下である。一度は買ってくれた顧客が離脱する理由を購買データなどのデータマイニングから探り出すことは難しい。一方、企業ではコールセンターなどに寄せられる苦情や要望、質問などの顧客の声を蓄積している。これらの顧客の声を分析して顧客が離脱する理由を突き止めようとする試みが行われている。

日本ユニシスでは顧客の声に対応するために、テキストマイニングを行う MiningPro 21 文書マイニングシステム、顧客の声の管理と分析を行うための顧客の声システム CVPro を提供している。

1) テキストマイニングの必要性 (MiningPro 21 文書マイニングシステム)

ある出版社の雑誌に読者相談のページがある。そのページを切り取り上段に悩み事や相談事を書いて送ると下段に回答を書いて送り返してくれる。これは読者サービスであるがその出版社にとっては情報収集の一つの手段でもある。読者の悩み事や相談事からヒントを得て新しい商品を開発しようというものである。ところが読者相談の文書も蓄積されて量が多くなると体系的に読むことができなくなる。担当者ごとのテーマに沿った文書だけを抽出できないかという要求が発生する。これを切っ掛けに日本ユニシスではテキストマイニングの研究を開始し、MiningPro 21 文書マイニングシステムを開発し提供している。MiningPro 21 文書マイニングシステムでは、大量の文書を基準を与えずに内容が類似したものをグループ化する「文書分類」と、苦情か、要望かなどあらかじめ決められたカテゴリに文書を判別する「文書判別」を行うことができる。

2) 顧客の声はリスク

企業は顧客との関係強化のためにコールセンターを設置して、顧客からの苦情や要望を聞いている。顧客の声は前述の出版社のように新商品開発のヒントを得られる「宝の山」という考え方もあるが、顧客の声はリスクであると捉えることもできる。顧客の声には二つおりのリスクがある。一つは、オペレーショナルなリスクである。あるコンピュータ周辺機器メーカーでは1日に5千件もの問い合わせがある。これに対応するために数百席のコールセンターを設置しているが、人件費をはじめその維持費用は膨大になる。もう一つの

リスクは、顧客が離脱するリスクである。苦情、要望、質問などコールセンターに寄せられる顧客の声は顧客の不満の現れと考えることができる。不満があれば顧客は離脱する。顧客の声を分析して顧客の不満を解消することが必要になる。これらのリスクを解消するために、日本ユニシスでは顧客の声システム CVPro 対話管理系と CVPro 声の活用系を提供している。

3) オペレーショナルなリスク

オペレーショナルなリスクにもことおこりある。量が多いことによる費用の増大と不適切な対応による顧客の不満増大である。顧客の声への対応には、全部の声に対応しきれない、対応が遅くて手遅れになる、違う人の同じ質問に異なる回答をする、同じ人の同じ質問に同じ回答をする、同じことを何度言っても対応がないなどの問題がある。寄せられる声の数が多くなれば全部の声に対応しきれず、対応ができない顧客の不満が増大する。寄せられた声の管理が悪くいつまでも放置すれば手遅れになり顧客の不満は増大する。担当者ごとに判断が異なり同じ質問に異なる回答をすることはよくあることである。マニュアルどおりの回答をしていると同じ顧客の同じ質問に同じ回答をすることになる。顧客は前回の回答に満足しなかったために再度の質問をしているとすれば、顧客の怒りを買うことになる。顧客との対応の履歴が記録されていないとこのようなトラブルを起こす恐れがある。また、声の内容は一見重大には見えなくとも顧客にとっては深刻な問題であることもある。同じことを何度も訴えているにも係わらず聞いてもらえないケースもある。特に顧客とのチャネルが電話、葉書、電子メールなど複数ある場合、それぞれのチャネルごとに顧客の声を管理していると一人の顧客の声を全体を把握できなくなっていることがある。

これらの問題を解決するためには、先ず顧客の声の一元管理が必要である。電話、葉書、電子メールなどばらばらにきた声を一つのデータベースに格納して、同じ顧客からの声を一覧し対応の履歴がひと目で分かるようにしなければならない。回答や詫言状など対応が必要な声はその進捗を管理していつまでも放置されないようにしなければならない。回答作成も一から始めていたのでは効率が悪い。雛形を用意したり過去の回答を検索できるようにしておかなければならない。

4) CVPro 対話管理系によるオペレーショナルなリスクへの対応

CVPro 対話管理系は、顧客の声の①登録、②仕分、③振分、④回答、⑤返信の声の対応の一連の流れを管理するソフトウェアである。①登録は、寄せられた声を入力してデータベースに格納する機能である。当然電子メールなどはじめからデジタル化されているものは自動登録される。②仕分は、登録された声の苦情か要望か質問かなどの区分や内容による分類を行う機能である。区分付けや分類分けは担当者が声の内容を読んで行うのが普通であるが、文書マイニングの「文書判別」を適用することにより自動化することも可能である。ある企業ではコールセンターに寄せられる声を約 100 のカテゴリに分類しているが、文書マイニングを導入したことにより 7 割の工数を削減している。③振分は、回答を作成する部署に声の処理を依頼する機能である。顧客の声は多くの場合、お客様相談室というような専任部署に寄せられ、回答もそこで作成される。しかし、専門性の高い内容や是非とも担当者に読んでもらいたい声などは、担当部署に転送することができる。担当部署も文書マイニング「文書判別」で自動判定させることができる。④回答は、実際に回答文を作成する機能である。回答作成の前にはその顧客がどのような顧客かを知る必要があ

る．そのために顧客との過去の対応の一覧と取引の明細が表示されるようになっている．重要顧客といわゆるクレームと呼ばれる顧客では回答内容も異なる．回答文の作成を容易にするために雛形が用意されているが，雛形に無いような難しい内容には過去にベテラン担当者が作成した回答文を参考にしたいことがある．このとき顧客から寄せられた声そのものを検索文にして検索すると，キーワードが一致しなくても内容が近い過去の声をリストアップしてくれる機能がある．文書マイニングの応用の類似検索機能である．これにより内容が近い過去の声の回答文を参考にすることが簡単にできるようになる．⑤返信は，回答文に前文，末文などの挨拶文を付加し顧客に回答を返す機能である．葉書や封書で来た声は封書で，電子メールできた声は回答も電子メールで返すのが普通である．以上のように CVPro 対話管理系は寄せられた顧客の声を間違いなく処理するためのシステムである．声の仕分けや回答作成に文書マイニング技術を適用して担当者の負担を軽減する工夫がなされている．これによりオペレーショナルなリスクを回避することが目的のシステムである．

5) 顧客が離脱するリスク

ある化粧品会社では定期的に顧客の声と顧客の離脱の関係を調べている．前期に購買があり今期に購買がなくなった顧客を離脱したと考え，離脱前の顧客の声を調べてどのような声を発した顧客に離脱が多いかを調査している．顧客を累積購買回数によって新規顧客，一般顧客，優良顧客に分類する．購買 1, 2 回の顧客は製品の利用が習慣化していない新規顧客である．購買 3, 4 回の顧客は習慣化初期の段階にある一般顧客である．購買 5 回以上の顧客は利用が習慣化している優良顧客である．各顧客層ごとに離脱者が発した声を多い順に 3 件ずつあげると以下のものであった．新規顧客は，質問，苦情，製品不良に関する声が上位 3 位である．一般顧客は，安全性，商品復活，感謝に関する声である．優良顧客は，感謝，肌の悩み，ファンデーションに関する声である．新規顧客はその企業や製品に対して理解度が低く，質問や苦情などが多くなるとされる．最初の購買で質問，苦情などの不満が解消されなかったり，製品不良など不愉快な思いをしたりすれば直ぐに離脱することは容易に考えられる．一般顧客に商品復活があるが，せっかく習慣化し始めたところでリニューアルなどでその製品が販売中止になれば離脱しても仕方がない．優良顧客でも肌の悩みの解消やファンデーションの機能性など製品に対する期待が満たされなければ離脱する．また何度も購買を重ねているにもかかわらず，いつまでも感謝などの具体性のない意見を言う顧客はその企業に対するロイヤリティは低いと考えられる．このように顧客の声から顧客の離脱理由を推定することは可能である．このような分析結果を利用してある企業は新規顧客に対して特別なホームページを用意している．新規顧客に対し企業や製品の理解を促すための対策である．顧客の声を分析した結果，新規顧客の質問や苦情に応えようというものである．顧客の声を分析することにより，それぞれの顧客層に適した対策を講じることが可能になる．

上記の調査から半年後に同様の調査を行った結果，各顧客層の離脱者の声の内容が変ってきていることが判明した．それぞれの多い声は以下のとおりであった．新規顧客は，苦情，製品不良，質問が上位の三つである．一般顧客は，安全性，製品不良，商品復活である．優良顧客は，製品不良，感謝，肌の悩みである．半年前の調査では一般顧客と優良顧客の上位 3 位内には製品不良は入っていなかった．ある程度の製品不良はそれらの顧客に

は容認範囲であったものと思われる。それが半年後には全部の顧客層の上位に製品不良が入っている。その半年間は牛海綿状脳症（BSE）の牛が発見されたり、新聞に毎日のように製品不良、異物混入、消費期限超過、不正原材料などのニュースや製品回収の広告が出ていた頃である。多少の製品不良などには頓着しない人も毎日のように製品不良や製品回収のような報道が繰り返されれば、自分が愛用する製品にも疑いを持つようになる。このように企業はいつものように変わりなく製品を作り続けサービスを提供していても、顧客の方の意識が変わっていることはよくある。これに気付かず旧態依然の対応を続けていれば多くの顧客が離脱する危険がある。顧客の声を監視して顧客の変化を見抜くことが必要になる。

6) CVPro 声の活用系による顧客が離脱するリスクへの対応

CVPro 声の活用系は、速報、報告、分析の三つの機能を持つ。前述のように離脱した顧客の声を分析すれば、どのような顧客がどのような理由で離脱するかを推定することができる。CVPro 声の活用系「分析」機能は、顧客の声と購買の関係を分析する機能である。これにより顧客が離脱する理由を推定することができる。顧客の意識や行動は変化するので定期的に分析を行う必要があるが、毎日それを行うのは困難である。そこで速報機能により顧客の声を監視することが必要になる。CVPro 声の活用系「速報」機能は、毎日寄せられる声を自動的に分類・集計して速報画面に表示する機能である。件数に大きな変化があれば警告が発せられるので、顧客の変化にいち早く気付くことができる。分析や速報により問題が発見されたら、報告機能により詳細な分析を行うことになる。CVPro 声の活用系「報告」機能は、自由な軸で顧客の声を集計し、ドリルダウンすることにより詳細な原因を追究することができる。

顧客の声を毎日人手で分析することは困難である。基本的には全部の声を読まなければならないので担当者にとっては時間も掛かり苦痛を伴う作業である。日に数件の声ならばそれも可能であろうが数が多くなればシステム化による支援が必要である。CVPro では、分析機能により顧客の基本的な特性を把握しておく。それが変化していないかを速報機能で監視する。問題が見付かれれば報告機能で原因を追究する。CVPro 声の活用系はできるだけ人手を排除して顧客が離脱するリスクを回避するための機能である。

ある企業は、顧客の声システム導入の効果を三つあげている。1) 先ず顧客の変化が分かるようになった、2) その結果リスク管理ができるようになった、3) また従業員の意識やスキルの向上に役立った、である。従業員が主観ではなく客観的に物事を見るようになり、考える力が付き文章力も判断力も向上したという。また、顧客の声システム CVPro はリスク発見ツールであるという。以上のように、CVPro は、顧客の声によるオペレーショナルなリスクと顧客が離脱するリスクの二つのリスクを軽減するためのシステムである。

5. おわりに

日本ユニシスは、10年以上にわたってマーケティング関連の分析システムとして、データベースマーケティング・システムの IMPACT DM/MA、データマイニング・ソリューションの MiningPro 21、顧客の声システムの CVPro を提供してきた。それぞれのソフトウェアの狙いは明確である。IMPACT DM/MA は戦略面および販売面における顧客情報の活用による顧客価値の増大である。MiningPro 21 は初心者では難しい分析モデルも同時に提供しデータマ

イニング導入の障壁を下げようというものである。CVPro は人手では苦痛を伴う顧客の声の分析を自動化して声に関連するリスクの解消である。このようにそれぞれの目的がはっきりしているところに日本ユニシスの分析系ソフトウェアの特徴がある。これらは単なるツールの提供だけではなく、導入時や運用時のコンサルティングを行うことにより、より大きな効果を上げている。導入前にはテスト分析を行い、導入の効果と投資の採算を事前に検証することによってその企業に最適なシステムを提案している。導入後も運用の支援や新たな課題に対するコンサルティングを行い陳腐化することのないよう常に最新の技術を提供している。このようなコンサルティングによって得られた新たな知識は次の開発に反映され、ソフトウェア自身もより実務的なツールに成長している。日本ユニシスの分析系ソフトウェアはさまざまな企業の知識や経験が集約されたものであるといえる。

- 参考文献**
- [1] Hughes, A., Strategic Database Marketing, McGraw Hill, 1994. (秋山耕監訳, 小西圭介訳, 『顧客生涯価値のデータベース・マーケティング』, ダイアモンド社, 1999)
 - [2] Blattberg, R., Getz, G. & Thomas, J., Customer Equity : Building and Managing Relationship as Valuable Assets, Harvard Business School Press, 2001.
 - [3] 原田保, 松田芳雄, 他, 『カスタマーマイニング』, 日科技連出版社, 2003.
 - [4] 松田芳雄, 「顧客分析とデータマイニングの動向」『ユニシス技報』68号, 日本ユニシス, 2001.
 - [5] Rud, O., Data Mining Cookbook, John Wiley & Sons, 2001.
 - [6] 松田芳雄, 「報告書」日本ダイレクトマーケティング学会自主研究プロジェクト, 自主研究プロジェクト報告書『顧客情報の活用実態と課題(データマイニングの適用を含めて)』, 日本ダイレクトマーケティング学会, NO. 2004 00001, 2004.
 - [7] 香田正人, 中川雄市, ルディー和子, 千葉潔, 『顧客情報の利用実態と課題(データマイニングの適用を含めて)』, 日本ダイレクトマーケティング学会第三回全国研究発表大会報告要旨, 日本ダイレクトマーケティング学会, 2004.
 - [8] 林田英雄, 脇森浩志, 「テキストマイニング技術とその応用」『ユニシス技報』84号, 日本ユニシス, 2005.
 - [9] 田代重信, 三池良洋, 「顧客の声システム CVPro」『ユニシス技報』84号, 日本ユニシス, 2005.

執筆者紹介 松田芳雄 (Yoshio Matsuda)

1974年慶応義塾大学工学部管理工学科卒業。同年日本ユニシス(株)入社。オペレーションズ・リサーチ, 統計解析関係のシステム開発に従事。現在, 日本ユニシス・ソリューション(株)に所属。日本オペレーションズ・リサーチ学会, 日本シミュレーション学会, 情報処理学会会員。