

2000年8月31日

SUGI-J 2000

# Webログを活用したデータマイニング

柳澤 彩子 松本 雅彦  
株式会社CSK  
技術企画 推進本部

## Agenda

- データマイニング成功のポイント
- Web分析の問題点
- 事例紹介
- Webデータ活用のポイント

## 分析事例

- 第三者不正利用検知
- 長期延滞者予測
- 退会者予測
- 顧客分析
- キャンペーン分析
- FSP利用促進分析
- チャネル分析
- アンケート分析

## データマイニング成功ポイント

- 現状の問題・目的を整理
  - データマイニングの適用部分を定義
- データの準備の確認
  - 件数、質のチェック
  - 必要な情報を収集する手段、仕組みの検討
- データ加工
  - データ統合
  - データの加工、選択がモデル精度に依存

## Webデータの活用

### ■ アクセスログサンプル

```
XXX.XXX.XXX.145 - - [25/May/2000:16:20:07 +0900] "GET  
/vendor/sun/ HTTP/1.0" 200 1954 "http://www.tdl.csk.co.jp/vendor/"  
"Mozilla/4.7 [ja] (WinNT; I)"  
XXX.XXX.XXX.101 - - [25/May/2000:16:26:07 +0900] "GET /  
HTTP/1.0" 304 - "-" "Mozilla/4.5 [ja] (Win98; I)"
```

## Web分析 現状における問題点

- ログを取っていない。標準ログしかない。
- 大量のログをどう扱っているかわからない。
- ログと顧客IDをマージできない。
- データが十分集まっていない。
- その他

## 事例 概要

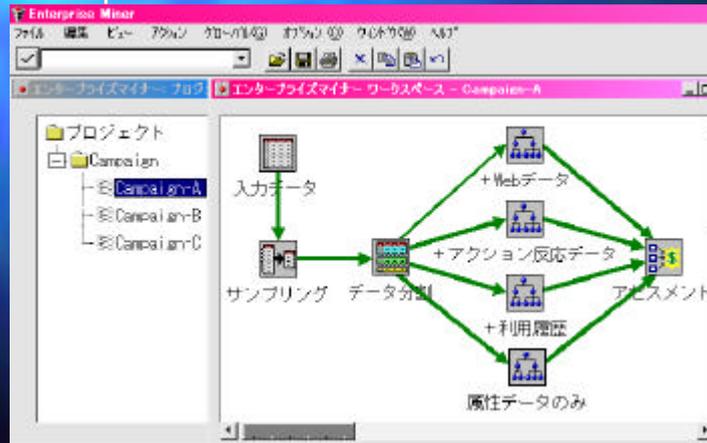
- ROIの高いキャンペーンを行うため  
過去に実施したキャンペーンのデータなどから反応率の高い顧客を抽出するモデルを作成する  
Webデータを用いて精度が上がるのか？

## 事例 使用データ

- 属性データ(性別・年齢・職業・住所・家族数)
- 取引履歴(金額集計・回数集計・取引期間)
- 反応データ(レスポンスデータ・サービス利用有無)
- 他モデルスコア
- Webデータ
  - インターネット購入有無、アンケート回答データ、メール配信サービス利用有無、特定ページのアクセス数

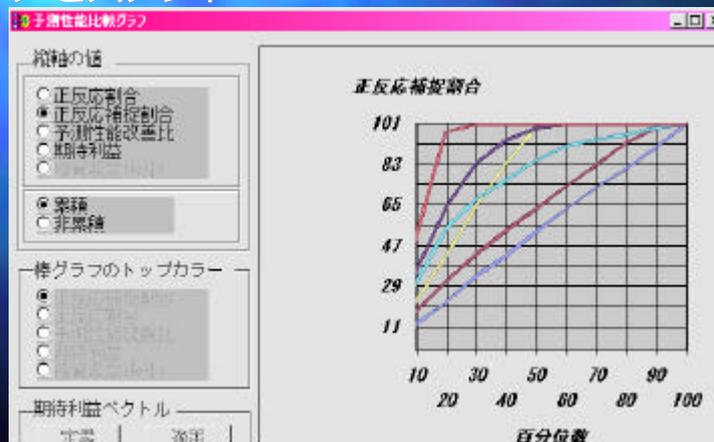
# 事例 : プロジェクトフロー

## ■ Enterprise Miner



# 事例 結果

## ■ アセスメント



## + Webログの相乗効果

### ■ 顧客開拓のSTEP

### ■ 顧客維持のSTEP

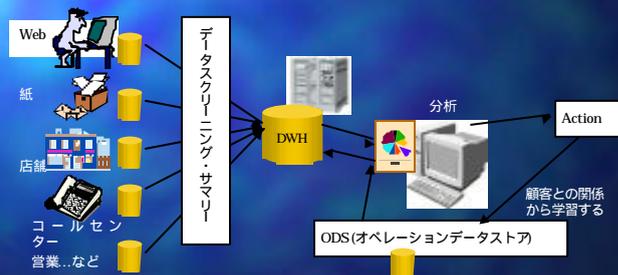


## 顧客獲得のSTEP

- アクセス数、時間
  - パフォーマンスの問題改善、キャパシティ計画
- 人気ページ、クリックストリーム
  - デザイン、サービスの再構築
- アクセス経路
  - バナー広告の最適化
- リピーター、好み
  - オプトインメール活用

## 顧客維持のSTEP

- 業務全体のサイクルを意識した仕組み
- サイクルを回して分析精度、コミュニケーション向上



## まとめ

