

# 模糊聚類演算法於 CRM 客戶分類之研究

## A Study of Fuzzy Clustering Algorithm for CRM Customers Classification

劉家驊 歐光榮

醒吾技術學院資訊科技應用研究所  
961620013@mail.hwc.edu.tw

### 摘要

準確的客戶分類是現代企業經營有效地實施客戶關係管理的基礎。客戶分類是根據客戶屬性來劃分客戶的集合，透過獲得的客戶類別來分析和預測客戶的消費模式，建立起一對一的客戶服務體系，以利企業實行進一步差異化客戶的管理。由於客戶分類問題涉及的因素眾多，許多因素帶有模糊性，且分類的標準常因為分類目的不同而有差異，因此沒有一種通用的方法適合各種客戶分類問題。

本研究提出 CRM 系統基於模糊 ISODATA 聚類法的客戶分類模型，利用選取客戶分類的特徵指標，並且引入了模糊群聚(Soft Clustering)隸屬度的概念，建立可以給出客戶各類分辨群組的隸屬度(是可以屬於不同的客戶群)，相較於以往硬式聚類(Hard Clustering)單一隸屬度，能更好地表現出客戶特徵。

研究中先對模糊 ISODATA 聚類演算法中初始劃分矩陣和分類數的確定，並使用最大矩陣元素法，求得最佳化演算法中其他參數的值。最後在 CRM 系統實際驗證和分析中，以改進的模糊 ISODATA 聚類演算法對汽車銷售公司實施客戶模組分類，所得到的聚類結果可幫助行銷人員做出預測，制定出針對客戶差異化的行銷策略，提高客戶服務的貢獻價值。

關鍵字：模糊聚類演算法，最大矩陣元素法，CRM，客戶分類

### 壹、緒論

客戶關係管理 (Customer Relationship Management, CRM) 旨在改善企業與客戶之間關係的新型管理機制，其主要精神係以顧客為核心發展各類行銷策略與活動，來強化與顧客之間關係，提升或延續顧客價值。因為就企業永續經營角度而言，企業的價值取決於所擁有的顧客數量與品質，當所獲得有價值的顧客越多，企業獲利的機會也就

越高。不同區隔的顧客，所能創造的利潤大小亦會有所不同，在有限的資源分配下，企業如何利用相關的指標，評估顧客能為企業帶來的價值為何，據此篩選出真正能創造利潤的顧客(消費力強或忠誠度高)，做為主要的目標客群(Target Customer)，並藉由顧客特性分析，適時、精準地投注資源，方能經濟、有效地達成顧客關係管理，創造最有效率的獲利[1]。

資料探勘(Data Mining, DM)在近代行銷領域中，佔有舉足輕重的角色，其使用範圍亦廣，目前主要應用為顧客關係管理方面。以國內某汽車公司為例，即藉由資料倉儲、資料探勘技術，針對顧客的商業行為進行分析，作為顧客關係管理重要參考依據。企業經由顧客交易紀錄的蒐集、記錄，便能依據歷史資料，分析出顧客決策的消費模式，預測對於各項行銷活動之可能反應，並可延伸推測資料庫外，具有類似行為模式顧客之可能回應，以進行新客源開發，擴大其應用範圍。資料探勘不僅可以發揮與行銷研究相類似之功能，更能持續不斷的監控顧客行為之變化，有效掌握市場變化脈動，隨時調整因應策略[2][3]。

由於目前較常用的聚類分析法(Cluster K-Means Analysis)，應用在某汽車銷售公司顧客分類判定，並不能精準地且完備做出分類，因此本研究擬以不同概念的模糊聚類演算法 (Fuzzy Clustering Method) 來佐以詮釋，兩者一併應用互相補其不足之處，期能更準確地將消費者資料做有系統整理及分類。由於影響消費者購買車輛的因素眾多，除對品牌有強烈忠誠度的客戶外，一般消費者監測資料的不確定性頗高。本研究先就現有 CRM 系統產生客戶分佈的聚類特性，希望能夠針對各資料的分佈情形，產生適合該種分佈的聚類中心(Cluster Center)，進而能找出較可靠的範圍畫定規則。再來比較原本的數值資料，依照監測資料的特性予以分群後，就可以得知問題群集(新的或將流失的客戶群)，配合即時行銷策略，開發或挽留特殊消費族群，在時間的管理成本會減少許多，建立起一對一的客戶服務體系，

實行差異化客戶管理，讓行銷決策者能正確、有效率地做出適當的處理及反應。

## 貳、文獻探討

### 一、客戶分類指標的選取

由於客戶分類問題涉及的因素眾多，許多因素帶有模糊性，並且分類的標準因為分類目的的不同而有所不同，因此沒有一種通用的方法適合各種客戶分類問題[4]。汽車銷售公司可根據客戶資料庫中已有的類型資訊的不同和自身管理的需要進行具體的分類。

客戶分類結果的正確與否取決於分類指標和分類方法的選擇。用於分類的指標要能反映客戶特徵以及銷售公司進行客戶管理的目的，分類指標的選擇需要遵循一定的客觀規律，並應根據銷售公司所處的行業特點以及公司本身的實際情況來選擇恰當的分類指標。

傳統上客戶分類的依據是客戶的統計學特徵（年齡、性別、收入、購買產品等），這種細分方法存在著很大的誤差。目前較常見的客戶分類方法有基於量化客戶價值的分類和基於指標組合的客戶分類方法，而指標組合的客戶分類常採用客戶利潤率、忠誠度和信用度的組合來分類客戶，基於客戶行為的客戶購買傾向分析是衡量客戶忠誠度的常用方法[5]。

在客戶行為分析中，RFM 指標是較為常用的評價指標。近度（Recency）衡量從最近一次購買到當前的時間。近度時間太長被廣泛認為是客戶行為已經發生變化的信號。研究發現，R 越大的客戶越有可能與企業達成新的交易。R 越大，企業保存的該客戶的資料就越準確，因為企業擁有的資料會迅速的失效，每隔一年約有 50% 的資訊變得不準確了。頻度（Frequency）衡量在一段給定的時間內客戶購買的次數，用以識別經常接觸企業的客戶。交易次數越多的客戶越有可能與企業達成新的交易。值度（Monetary）衡量購買的平均金錢價值。M 越大，越有可能再次回應企業的產品和服務。本文的汽車銷售 CRM 系統應用中，採用 RFM 作為分類衡量指標，以客戶的 RFM 行為作為對客戶忠誠度和客戶利潤率的類比衡量[6]。

### 二、模糊聚類演算法

模糊聚類相對於硬聚類（任一元素只能屬於某一類）能更好地體現客戶特徵[7]。因為在對客戶分類分析的時候，並不能指出某個客戶一定屬於或一定不屬於某一客戶群體。恰恰相反，一個客戶樣本可以屬於不同的客戶群。模糊聚類引入了隸屬度的概念，就很好地體現了這一事實。如果採用硬聚類

就不能反映出客戶在屬性方面的特徵。而模糊聚類則可以給出客戶分別屬於各類的隸屬度，從而幫助行銷人員制定出針對客戶的行銷策略，以提高客戶的價值貢獻。

模糊 ISODATA（Interactive Self-Organizing Data Analysis Technique）聚類法是基於模糊劃分的思想，利用反覆運算方法，在泛函極值意義下，不斷修正聚類中心的局部優化演算法[8]。

### 三、模糊 ISODATA 法

設有  $n$  個樣本的總體（論域） $X = \{X_1, X_2, \dots, X_n\}$ ，每一個樣本有  $s$  個特徵

$X_j = (X_{j1}, X_{j2}, \dots, X_{js})$  ( $j=1, 2, \dots, n$ )

首先，取定分類數  $c$ ，軟劃分矩陣  $U = [u_{ij}]_{c \times n}$  的全體構成軟劃分空間）

$M_{fc} = \{U_{c \times n} | U_{ij} \in [0, 1], \forall i, j\}$ ;

$$\sum_{i=1}^c u_{ij} = 1, \forall j;$$

$$0 < \sum_{i=1}^c u_{ij} < n, \forall i (i=1, 2, \dots, c; j=1, 2, \dots, n)$$

其中： $u_{ij}$  表示第  $j$  個樣本  $X_j$  隸屬於第  $i$  類的隸屬度。

對  $U$  需滿足以下三個條件：

1. 每個元素  $u_{ij}$  在  $[0, 1]$  上取值；
2. 每列元素之和為 1，即同列中分屬於各類的隸屬函數之和為 1；
3. 每行元素之和大於 0，即每類中至少有一個樣本。

定義聚類中心： $V = [V_1, V_2, \dots, V_c]^T$ ，其中  $V_i = (v_{i1}, v_{i2}, \dots, v_{is})$ ，( $i=1, 2, \dots, c$ ) [4]。第  $i$  類的中心  $V_i$  即人為假想的理想樣本，它對應的  $s$  個特徵值是該類樣本所對應的特徵值的平均值。

構造目標函數：

$$J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|X_j - V_i\|^2 (m \geq 1)$$

其中  $\|X_j - V_i\|$  表示第  $j$  個樣本與第  $i$  類聚類中心之間的距離； $J_m(U, V)$  表示所有待聚類樣本與所屬類的聚類中心之間距離的平方和。

參數  $m$  是加權指數，控制著模糊類間的分離程度；Bezdek 在文獻[9]中對參數  $m$  的確定進行了模擬試驗研究，他的試驗結果表明，參數  $m$  以採用 2 為優。

為了得到最佳劃分矩陣  $U^*$  對應的聚類中心  $V^*$ ，從而確定最佳分類結果，就必須使  $J_m(U, V)$  達到最小，即  $J_m(U^*, V^*) = \min_{U \in M_{fc}} \{J_m(U, V)\}$ 。

一般而言，上述目標函數的極值求解相當困難，但 Bezdek 已經證明[10]) 當  $m \geq 1$ ， $X_j \neq V_i$  時，可以透過下面的方式進行反覆運算，並且運算過程是收斂的，這就是模糊 ISODATA 方法。其步驟為)

1. 選定分類數：  
 $c$  ( $2 \leq c \leq n-1$ )，取一初始模糊劃分矩陣  $U^0$ ，逐步反覆運算 ( $l=0, 1, 2, \dots$ )。

2. 計算聚類中心特徵陣：

$$V^{(l)} = (V_1^{(l)}, V_2^{(l)}, \dots, V_c^{(l)})^T$$

$$\text{式中， } V_i^{(l)} = \frac{\sum_{j=1}^n (u_{ij}^l)^m X_j}{\sum_{j=1}^n (u_{ij}^l)^m}$$

3. 修正劃分矩陣  $U^{(l)}$ ：

$$u_{ij}^{(l+1)} = \frac{1}{\sum_{k=1}^c \left( \frac{\|Xi - V_i^l\|}{\|Xi - V_k^l\|} \right)^{\frac{2}{m-1}}} \quad (i=1, 2, \dots, c; j=1, 2, \dots, n)$$

4. 用矩陣範數：

$\|\cdot\|$  比較  $U^{(l)}$  與  $U^{(l+1)}$ 。對取定的  $\varepsilon > 0$ ，若  $\|U^{(l+1)} - U^{(l)}\| < \varepsilon$ ，則反覆運算停止，取  $U = U^{(l+1)}$  作為最終劃分矩陣；否則  $l=l+1$  轉(2)繼續反覆運算。

應用以上演算法得到模糊劃分矩陣  $U^{(l+1)}$  和聚類中心  $V^{(l)}$  是相對於分類數  $c$ 、初始模糊劃分矩陣  $U^0$ ，加權指數  $m$  和收斂精度  $\varepsilon$  的最佳解。

如果改變  $c$ 、 $U^0$ 、 $m$  和  $\varepsilon$ ，則可得到許多局部最優解。若從這些最優解中選出最佳，則需要有鑒別模糊 ISODATA 聚類效果的指標。

鑒別聚類效果可用下列指標：

1. 分類係數  $F_c(U)$ ：

$$F_c(u) = \frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c u_{ij}^2$$

$F_c(U)$  越接近 1，最終分類的模糊性越小，聚類效果越好。

2. 平均模糊熵  $H_c(U)$ ：

$$H_c(u) = -\frac{1}{n} \sum_{j=1}^n \sum_{i=1}^c u_{ij}^2 \ln u_{ij}$$

$H_c(U)$  越接近於 0，聚類效果越好。

分類完成後，對於新樣本類別歸屬的判斷遵循原則如下[11]：

原則 1：

設最後求得的聚類中心向量為：

$$V = (V_1^*, V_2^*, \dots, V_c^*)^T$$

$\forall x_k \in X$ ，

如果： $\|x_k - V_c^*\| = \min_{1 \leq j \leq c} (\|x_k - V_j^*\|)$

則將樣本  $x_k$  歸於第  $i$  類。

該判定原則的意義是：新樣本  $x_k$  與哪一聚類中心最接近，就將它歸到那一類。

原則 2：

設最後求得的模糊劃分矩陣為：

$$U^* = (u_{ij}^*), \quad \forall x_k \in X, \text{ 在 } U^* \text{ 的第 } k \text{ 列中, 如}$$

果  $u_{ik}^* = \max_{1 \leq j \leq c} (u_{jk}^*)$ ，則將樣本歸於第  $i$  類。

該判定原則的意義是：新樣本  $x_k$  對哪一類 ( $X$  上的模糊子集) 隸屬度最大，就將它歸到那一類。

#### 四、確定分類數和初始劃分矩陣

模糊 ISODATA 聚類演算法是一種效果理想的聚類分析方法，只要根據研究物件的具體情況確定預分類數  $c$  就可以實現對樣本的最優劃分。然而，由於系統的高度複雜性和指標的模糊性，常常是事先無法確定預分類數。即使對系統有一些瞭解，事先給定預分數，而實際上是對演算法的一種人工幹預，很可能傷害了分類的科學性。

初始模糊劃分矩陣  $U^0$  的選取對聚類結果影響不大，但  $U^0$  元素的取值愈接近各樣本在不同類別的隸屬度，反覆運算收斂愈快，分類也愈精確，聚類中心值也愈接近實際情況。

為了得到最優分類，應確定合理的分類數  $c$  ( $2 \leq c \leq n-1$ )。採用模糊聚類最大矩陣元素法確定分類數) 首先構造模糊相似矩陣，矩陣的元素可由下面的數量積公式求出)

$$r_{ij} = \begin{cases} 1 & i = j \\ \frac{\sum_{k=1}^n x_{ik} x_{jk}}{\max_{i,j} \sum_{k=1}^n x_{ik} x_{jk}} & i \neq j \end{cases} \quad (i, j = 1, 2, \dots, n)$$

模糊相似矩陣任一行的最大矩陣元素 (主對角線除外) 與截集水準  $\lambda$  相等，用最大矩陣元素代替截集水準  $\lambda$ ，把  $\lambda$  值從大到小排列，使值相近的樣本子集聚為一類，根據  $\lambda$  值的突躍區將樣本分為  $c$  類，作為模糊綜合聚類反覆運算法的初始分類數，據此設一初始劃分矩陣來求最優分類。

#### 五、模糊聚類法其他參數的探討

1. 加權指數  $m$ ：

加權指數  $m$  控制著模糊類間的分類程度， $m$  值的選取對整個聚類過程和聚類結果有較大影響。參數  $m$  越接近 1，分類的模糊性越小，當  $m = 1$  時，分類變成硬分類；參數  $m$  越大，分類的模糊性越大，它的意義也更不明確。由於  $m$  出現在泛函  $J(U, V)$  中作為一個指數，它的值不宜太大，否則會引起失真。另外由於  $m-1$  在修正劃分矩陣的運算式中作為分母，故  $m$  值又不能太接近於 1，否則

會引起計算溢出。試驗發現，m 值越小，反覆運算次數越少，分類速度越快，劃分矩陣 U 的值越趨向於 0，1 兩極，聚類效果較好。

## 2. 距離類型[12]：

$$\text{目標函數 } J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m \|X_j - V_i\|^2$$

中需要計算第 j 個樣本與第 i 類聚類中心之間的距離。在應用中最常用的距離有)

### (1) 最大值距離：

$$d_1(\|X_j - V_i\|) = \max_{i \leq k \leq s} |x_{jk} - v_{ik}|$$

### (2) 歐氏距離：

$$d_2(\|X_j - V_i\|) = \sqrt{\sum_{k=1}^s (x_{jk} - v_{ik})^2};$$

### (3) 絕對值距離：

$$d_3(\|X_j - V_i\|) = \sum_{k=1}^s |x_{jk} - v_{ik}|^2;$$

### (4) 閔可夫斯基距離：

$$d_4(\|X_j - V_i\|) = \left( \sum_{k=1}^s |x_{jk} - v_{ik}|^p \right)^{1/p}$$

從實例運行結果看來，採用不同的距離類型對聚類結果影響不大，但仍有差異。用閔可夫斯基距離效果最好，歐氏距離次之，其餘的較差。在採用閔可夫斯基距離時，參數 p 的取值選擇在 11 到 13 之間，從距離算式看出，p 太大會引起溢出。

值得注意的是，模糊 ISODATA 聚類演算法是基於距離的演算法，這樣聚類結果受資料集中孤立點的影響很大。由於這個演算法不能識別出孤立點，所以只要存在一個或以上孤立點，聚類中心的位置就會偏移。

## 3. 收斂精度 $\varepsilon$ ：

$\varepsilon$  的取值是精度要求，對於整個聚類結果影響不大， $\varepsilon$  取值越小，精度越高，聚類效果越好。試驗表明，為了保證聚類的可靠性， $\varepsilon$  取值小於  $10^{-4}$ 。

## 叁、研究方法

在汽車銷售 CRM 系統中，有了最大矩陣元法和模糊 ISODATA 綜合的模糊聚類方法，就可對購買汽車的客戶樣本進行分類了。但在這之前需要對樣本的特徵值資料進行預處理)

考慮到各影響因素量綱的不同以及各因素的影響程度的不同，需對原始資料進行歸一化權重處理，採用下式)

$$x_{ij} = \frac{x_{ij} - x_{j\min}}{x_{j\max} - x_{j\min}} W_j$$

式中， $x_{j\min} = \min\{x_{1j}, x_{2j}, \dots, x_{nj}\}$ ； $x_{j\max} = \max\{x_{1j}, x_{2j}, \dots, x_{nj}\}$ ； $W_j$  為樣本的第 j 個影響因素的權重。在購買汽車的客戶分類模型中，考慮到 RFM 各指標對客戶行為的重要性相當，故權重取為等權重，即  $W_j = 1/s$  ( $j = 1, 2, \dots, s$ )。

下面給出了模糊綜合聚類演算法的代碼。本汽車銷售 CRM 系統是基於 B/S 結構實現的，採用 Microsoft SQL Server 資料庫，在 ASP 環境下開發，伺服器端腳本用 VBScript 編寫。

Dim m 加權指數

Dim dbIE 精度  $\varepsilon$

Dim ArrX () 樣本特徵陣

Dim ArrU0 () 初始劃分矩陣

Dim ArrU1 () 劃分矩陣

Dim ArrV () 聚類中心陣

Dim n 購買汽車的客戶的樣本數

Dim s 特徵數

基於模糊 ISODATA 的綜合聚類的主要步驟

Call pretreatment () 調用過程，對樣本原始資料進行預處理

ReDim ArrX (n, s)

Call computeC ()

調用過程，計算相似關係矩陣，根據最大矩陣元法，得到分類數 c

ReDim ArrU0 (c, n)

ReDim ArrU1 (c, n)

ReDim ArrV (c, n)

Call U0 () 調用過程，得到初始劃分陣

Do

Call calculateV () 計算聚類中心陣

Call correctU () 修正劃分矩陣

Loop Until compareU (ArrU1, ArrU0) < dbIE

反覆運算計算劃分矩陣，直到目標函數小於給定精度函數 compareU (Arr1, Arr2)，比較矩陣範數

Call evaluate (ArrU1)

計算分類係數和平均模糊熵，輸出以評價聚類效果下麵給出了修正劃分矩陣過程的詳細代碼

過程 correctU () 無傳入參數，修改總體變數 ArrU1

Sub correctU ()

Dim i 矩陣行數

Dim j 矩陣列數

Dim sum 求和

Dim k

For i = 1 To c

For j = 1 To n

```

sum = 0
For k = 1 To n
    累加求和，其中函數 disc ( i, Arr1, Arr2) 計
    算向量距離
    sum=sum+ ( disc ( 2, ArrX ( j) , ArrV ( i) )
    /disc ( 2, ArrX ( j) , ArrV ( k) ) ) (2/(m-1))
Next
ArrU1 ( i, j) =1/sum 的劃分矩陣
Next
Next
End Sub
函數 disc ( i, Arr1, Arr2) ) 計算向量 Arr1
與 Arr2 距離，整數 i 表示
距離類型 (取值 1, 2, 3, 4)
Function disc ( i, Arr1, Arr2)
'...
End Function

```

#### 肆、範例驗證

本研究中的客戶分類模型實際上已經在某汽車銷售公司已實施的 CRM 系統中應用。本節內容亦加以驗證說明，使用的實例資料是隨機抽取的 50 個購買汽車的客戶樣本。按照上述客戶分類模型，在對初始資料標準化後，求得模糊相似關係矩陣，根據最大矩陣元素法，得到分類數  $c = 5$ 。

在初始劃分矩陣上，應用模糊 ISODATA 方法，反覆運算計算劃分矩陣和聚類中心矩陣，其中參數  $m = 2$ ， $\epsilon = 0.00001$ ，距離類型採用歐氏距離。聚類結果如表 1。

表 1：最終聚類中心陣和模糊 ISODATA 的檢驗指標

聚類中心陣 (標準化指標)			檢驗指標	
			分類係數 $F_c(U)$	平均 模糊熵 $H_c(U)$
0.05286	0.22300	0.25967	0.942751	0.071643
0.11245	0.15438	0.18617		
0.09971	0.11385	0.11290		
0.15048	0.13462	0.17140		
0.19056	0.03692	0.03571		

最終結果得到的劃分矩陣  $U$  體現了購買汽車的客戶樣本對不同客戶類別的隸屬度，取兩個客戶樣本作分析。

樣本 1:  $U1 = (0.0013, 0.9263, 0.0722, 0.0002, 0.0000)^T$

樣本 2:  $U2 = (0.1475, 0.8451, 0.0001, 0.0056, 0.0017)^T$

劃分矩陣的每列的資料代表了該樣本對某類的隸屬度，每列之和為 1，也就是對各類的隸屬度為 1。根據樣本判定的擇近原則，樣本 1 屬於第 2 類，樣本 2 也屬於第 2 類。但從隸屬度看來，樣本 2 還具有類別 1 的特徵。

此驗證實例說明瞭客戶分類的複雜性，在對客戶 2 進行行銷活動時，不僅要按第 2 類客戶類別的特徵制定行銷策略，還要考慮它具有第 1 類客戶類別的特徵。

本分類模型劃分了不同客戶種類，區分了不同客戶種類的特徵。表 2 顯示對實例結果的 5 類客戶的特徵進行的分析結果。

表 2：對實例特徵分析的結果

類別	客戶數量	類別客戶	特徵分析
I	9	成熟客戶	頻繁與汽車銷售公司接觸，且累計交易量遠大於其他客戶
II	27	主要客戶	與汽車銷售公司接觸的時間較近，購買金額較大，但接觸的次數較少
III	6	新客戶	與汽車銷售公司接觸的時間較近，購買金額和接觸的次數都較少
IV	5	衰退客戶	較長時間沒有與汽車銷售公司接觸，然而其與汽車銷售公司的接觸頻度和價值貢獻都比較高，可能是有流失危險的有價值客戶
V	3	無價值客戶	較長時間沒有與汽車銷售公司接觸，且價值不大甚至無利潤的客戶

根據分析結果顯示，由於採用 RFM 客戶行為為分析指標作為購買汽車的客戶分類的指標，劃分的購買客戶類別的特徵驗證了客戶對汽車銷售公司的營

業利潤的貢獻率，並且提供分析客戶的忠誠度和客戶流失的可能性。

### 伍、結論與未來研究方向

準確的客戶分類是現代企業經營有效地實施客戶關係管理的基礎。模糊聚類分析法是非監督式識別的一個重要分支，在眾多領域已得到了廣泛應用，是以研究模糊聚類演算法及如何實際應用於現代 CRM 系統中的具有十分重要的意義。本文探討了模糊理論的基礎知識和模糊聚類的理論發展，深入分析模糊 ISODATA 演算法的觀念、特點和問題。在此基礎上，利用最大矩陣元素法確定合理的分類數和初始的劃分中心矩陣，解決了應用模糊 ISODATA 演算法由於要事先給出參數，而導致結果可能帶有主觀臆測性的問題，研究透過理論分析和反覆試驗比較聚類有效函數值，最佳化了模糊 ISODATA 演算法中其他參數的取值，除了在理論上更趨嚴謹，減小了演算法對初值的敏感性外，而最終聚類結果亦更趨穩定。

本研究先對 CRM 系統的發展與問題加以定義，並分析 CRM 系統的功能結構、實務發展和對企業帶來的衝擊。針對不易解決的客戶分類問題，進行了理論和概念上的探討，其次分析細分客戶的問題特性，探討現行客戶分類的方法和用途，為了達到有效客戶分類的目的，及為對待客戶服務差異化提供科學依據和基礎，已透過有針對性的顧客制定行銷策略，期以降低成本，並藉以提高企業的經濟效益。經由實證分析，只要有一定數量的客戶資料，就可應用本分類模型對客戶進行分類評價，找出每一類客戶的特徵，為企業提供了客戶分析的量化依據，為決策策略的正確實施提供了科學、客觀、可靠的資料保證。

研究內容，雖然實現了 CRM 系統中的客戶分類的要求，並對模糊 ISODATA 演算法進行了改進，但還有需要進一步改善部分，未來研究的工作包括以下幾個方面：

1. 改進的模糊 ISODATA 演算法仍不適用於進行非凸形狀的聚類和大小相差很大的聚類。另外，演算法對雜訊和邊緣和孤立點很敏感。如何解決，需要再深入的研究。
2. 應用方面：經過實例的驗證，該客戶分類模型能適合不同企業用戶的對客戶分類的要求。但對於大量資料的運算，演算法效率仍需要進一步的改進，以便能更高效的工作。另一方面，對模糊聚類演算法在 CRM 系統中其他方面的合理應用，有待繼續進行。

### 參考文獻

- [1]安迅資訊 (NCR)，『整合企業經營策略與顧客關係管理』，電子化企業經理人報告-電子化顧客關係管理，第 5 期，頁 20-25，民 2000。
- [2]陳佳麟，『資料探勘應用於顧客關係管理之研究』，台北大學企業管理學系碩士論文，2002。
- [3]王媚慧，『應用資料探勘於顧客的行為分析-以半導體業為例』，淡江大學資訊工程系碩士論文，2004。
- [4]M.R. Anderberg, Cluster Analysis for Applications, Academic Press, New York, USA, 1973.
- [5]閔庭祥，『顧客關係管理系統之價值模型建構』，中央大學資訊管理研究所博士論文，2001。
- [6]楊清潭，『應用資料探勘技術於顧客價值分析之研究』，東吳大學資訊科學系碩士論文，2003。
- [7]佐和正敏，1989，模糊理論的基礎與應用，森北出版。
- [8]蘇建源，『模糊邏輯與資料探勘技術為基礎在顧客關係管理上之研究與應用』，南華大學資訊管理研究所碩士論文，2004。
- [9]BEZDEK JC. Physical interpretation of fuzzy ISODATA[M].IEEE Trans, SystemsMan, Cybern, 1976, SME - 6.
- [10]BEZDEK JC. Pattern Recognition with FuzzyObjective Function Algorithms [M].New York, Plenum, 1981, 65 - 93.
- [11]HSU TH. An application of fuzzy clustering in group positioning analysis[J].Proc Natl Sci, Counc ROC (C), 2000, 10 (2), 157 - 167.
- [12]D Fisher.Improving inference through conceptual clustering[J].In Proc.1987.AAAI Conf.Seattle.WA.July.1987.461-465