

## 摘要

顧客是企業一項非常重要的資產，而顧客的價值與重要性也是企業極需要了解的部份。近年來，商業智慧的發展促使顧客關係管理能更進一步被探討，運用科技的方法來分析企業與顧客間的買賣關係，使企業接近具有潛力的顧客。

為了篩選具有潛力之顧客群，本研究首先運用不同且簡易的顧客終身價值 (Customer lifetime value, CLV) 模型估算顧客價值，希望透過不同的模型有效地對顧客進行區隔，然後以 logistic 迴歸方法找出不同顧客分群之關鍵變數。最後經由商業智慧軟體透過視覺化的介面，明白告訴行銷管理人員該如何透過顧客交易行為的變化制定適當之管理決策。

本研究以 635 位經銷商交易資料庫為資料來源，萃取相關資料以進行各種顧客價值指標之計算，並且比較顧客終身價值中的一般價值模型、顧客維持模型，並將顧客終身價值以 K-means 進行顧客區隔，區隔成高、中、低價值經銷商。比較兩種 CLV 模型結果證實顧客維持模型較適合本研究之資料型態。另外，本研究透過不同 logistic 迴歸方法來探索顧客群體，研究結果發現多項式 logit 較次序 logistic 迴歸適合於本研究之預測分析，並且發現 3 項價值指標能解釋不同價值程度之經銷商，並發現中維持價值經銷商為適合行銷資源配置之經銷商，因此最後以時序群集方法分析中維持價值經銷商，結果發現可將分類為偏好促銷方案經銷商、偏好無促銷方案經銷商 A 及偏好無促銷方案經銷商 B，最後本研究針對此三類經銷商制定適合之管理策略，使其達到更高的利潤貢獻度。

**關鍵字：**顧客終身價值、Logistic 迴歸、資料探勘、顧客關係管理

# Abstract

A customer is a valuable asset, but how to assess the value of a customer has become an important issue recently, especially in the area of customer relationship management. However, the calculation of the customer lifetime value (CLV) becomes complicated with the effort and extra cost to deliver a better model. In the other way, data mining has been used a lot to explore the consumer's buying pattern and to generate useful models to predict buyer's behaviors. Therefore, this study tried to demonstrate a procedure to predict a customer's CLV by combining technologies of data mining and logistics regression.

This study adopted the AdventureWorks database provided by Microsoft SQL Server 2005 as the data source, which contains transaction records of 635 resellers to a fictitious bike company for three consecutive years. The data of the first two years was the training set to evaluate models, and the last year data was used to verify the model obtained in the training set.

Firstly, this study used the training dataset to estimate reseller's CLV with two basic models: the model of the general value and customer retention model. Then K-means analysis was used to divide resellers into three groups: high, medium, and low CLV resellers. After comparing the group assignments of the training dataset and the verification dataset, this study found customer retention model was the better model to predict resellers' CLV.

Next, the study conducted the logistic regression on 14 possible explanatory variables to explain the CLV group assignment. The results showed that the stepwise multinomial logit regression was better than the ordered logistic regression in explaining the dependent variables. Furthermore, the number of purchase items, the amount of purchase, and the promotion expense were found to be reliable indicators for the possible CLV groups. This study also found that the medium CLV group was the most potential target for marketing.

Finally, a data mining technology called sequence clustering was used to discover the association among purchase items toward the medium CLV group. Three clusters were identified: promotion-preferred resellers, non promotion-preferred resellers A, and non promotion-preferred resellers B. Marketing strategies were suggested to these three clusters for better profit returns.

**Keyword : Customer lifetime value 、 Logistic regression 、 Data mining 、 Customer management relationship**

# 目 次

ABSTRACT.....	V
目 次.....	VII
表目錄.....	VIII
圖目錄.....	IX

## 表目錄

表 2.1 顧客關係管理定義 .....	錯誤!	尚未定義書籤。
表 2.2 顧客終身價值之定義 .....	錯誤!	尚未定義書籤。
表 2.3 資料探勘技術與顧客關係管理上之應用 .....	錯誤!	尚未定義書籤。
表 2.4 顧客關係管理與顧客終身價值、資料探勘相關研究 .....	錯誤!	尚未定義書籤。
表 3.1 整理後資料屬性-經銷商交易資料表 .....	錯誤!	尚未定義書籤。
表 3.2 混淆矩陣表 .....	錯誤!	尚未定義書籤。
表 4.1 各經銷商之總淨利計算 .....	錯誤!	尚未定義書籤。
表 4.2 2001 年至 2004 年之美國貼現利率 .....	錯誤!	尚未定義書籤。
表 4.3 各經銷商之一般現在價值計算過程 .....	錯誤!	尚未定義書籤。
表 4.4 各經銷商每期之貢獻利潤 .....	錯誤!	尚未定義書籤。
表 4.5 每期(季)之維持率 .....	錯誤!	尚未定義書籤。
表 4.6 經銷商之顧客維持價值 .....	錯誤!	尚未定義書籤。
表 4.7 最後集群中心點 .....	錯誤!	尚未定義書籤。
表 4.8 基本模型價值個數 .....	錯誤!	尚未定義書籤。
表 4.9 最後集群中心點 .....	錯誤!	尚未定義書籤。
表 4.10 維持模型價值個數 .....	錯誤!	尚未定義書籤。
表 4.11 基本模型之訓練集與驗證集比對分析 .....	錯誤!	尚未定義書籤。
表 4.12 維持模型之訓練集與驗證集比對分析 .....	錯誤!	尚未定義書籤。
表 4.13 次序 logistic 迴歸與多項式 logit 適合度與參數估計值 .....	錯誤!	尚未定義書籤。
表 4.14 多項式 logit 概似比檢定表 .....	錯誤!	尚未定義書籤。
表 4.15 多項式 logit 參數估計值 .....	錯誤!	尚未定義書籤。
表 4.16 多項式 logit 模型分類表 .....	錯誤!	尚未定義書籤。
表 4.17 資料表篩選變數 .....	錯誤!	尚未定義書籤。
表 4.18 時序群集狀態轉換分析表 .....	錯誤!	尚未定義書籤。

## 圖目錄

圖 1.1 商業智慧趨勢週期圖 .....	錯誤! 尚未定義書籤。
圖 1.2 2008 年流通業商用軟體投資金額分配表 .....	錯誤! 尚未定義書籤。
圖 1.3 研究流程 .....	錯誤! 尚未定義書籤。
圖 2.1 顧客關係管理策略績效衡量指標 .....	錯誤! 尚未定義書籤。
圖 2.2 顧客權益行銷目標範圍 .....	錯誤! 尚未定義書籤。
圖 2.3 情境 1 現金流量時間點 .....	錯誤! 尚未定義書籤。
圖 2.4 情境 2a 現金流量時間點 .....	錯誤! 尚未定義書籤。
圖 2.5 情境 2b 現金流量時間點 .....	錯誤! 尚未定義書籤。
圖 2.6 情境 3 現金流量時間點 .....	錯誤! 尚未定義書籤。
圖 2.7 離散型利潤曲線 .....	錯誤! 尚未定義書籤。
圖 2.8 連續型利潤曲線 .....	錯誤! 尚未定義書籤。
圖 2.9 資料探勘之應用 .....	錯誤! 尚未定義書籤。
圖 2.10 顧客關係管理策略連結架構 .....	錯誤! 尚未定義書籤。
圖 3.1 本研究架構與流程 .....	錯誤! 尚未定義書籤。
圖 4.1 資料來源檢視設定 .....	錯誤! 尚未定義書籤。
圖 4.2 索引鍵、輸入與可預測變數設定 .....	錯誤! 尚未定義書籤。
圖 4.3 時序群集群集圖表 .....	錯誤! 尚未定義書籤。
圖 4.4 時序群集之群集圖表與陰影變數(促銷類型).....	錯誤! 尚未定義書籤。
圖 4.5 時序群集之群集設定檔 .....	錯誤! 尚未定義書籤。
圖 4.6 時序群集之群集特性 .....	錯誤! 尚未定義書籤。
圖 4.7 時序群集之群集辨識 .....	錯誤! 尚未定義書籤。
圖 4.8 時序群集之狀態轉換 .....	錯誤! 尚未定義書籤。

