Retail Assets Management in Indian Banking: Credit Marketing Model using Data Mining Approach

Dinabandhu Bag[1] & Asit Mohanty[2]

Abstract

With the increasing competition, decreasing customer loyalty and plethora of banking products, it has become essential in today's sluggish economy for banks to accurately position themselves. This paper presents a credit marketing model using an advanced data mining for acquisition segmentation and profit forecast. The average profitability (NPV) could increase with the use of this approach since the profit accruals vary across segments. Profit Forecast Models are built across each segment during the average life of a product. The model implementation on a Bank's test data results in the robust segment distribution in groups of 22%, 29%, 14% and 35%, which forms a good representation of the population. Segments demonstrate distinction across their demographic and credit behaviour profiles. The average profitability in our model scenario over a random scenario stands at Rs. 168 per account and Rs.29 per account for the top two segments respectively.

Keywords: Clustering, Segmentation, Forecast, Profitability

1.0 INTRODUCTION

Credit Marketing has come a long way in today's economy of hard-hitting competition and diminishing customer loyalty. A well defined credit marketing strategy should bring information driven approach to the marketing of credit to help identify and to implement robust approaches to the problem of targeting. This could also improve efficiency and provide roll out opportunities for greater benefits within the bank. With the increasing level of cut-throat competition, decreasing customer loyalty and the increasing commoditisation of banking products, it has become essential in today's sluggish economy for banks to predict the behaviour of their customers to position themselves accurately. This paper intends to highlight the issue

Received April 23, 2009; Revised September 4, 2009

1. Research Scholar, University of Mysore, Mysore, email: dinabandhu.bag@gmail.com
2. Associate Professor, Xavier Institute of Management, Bhubaneswar, email: asit.mohanty@ximb.ac.in

of credit marketing in banks and proposes to adopt a profitability based approach to credit markets using an advanced data mining method. Banks must understand the need based behaviours of their customer prospects to target them with the suitable product offerings. Not all customers are created equal, nor do they represent the same value to the Bank. Nearly two-thirds of a bank's revenue comes from less than one-third of the customer base (Graeme and John, 2001; Tina, 2000). Banks attempt to target customer prospects with offers which should result in both acceptance and profit to the bank. The typical Banks' market segmentation involves the practice of heuristic segmentation to define markets. This means banks select prospects using deterministic splits, such as geography, income or turnover code, occupation code, etc and or a combination of them, which is referred to as the BAU (Business As Usual) approach. It is targeting of customers by means of a heuristic and qualitative based differentiation of customer value. The irony of this approach is that the definitions of the target segments are based on a combination of deterministic splits, which is modified and altered by the marketing manager, every time, when there is a new campaign or a new offer. Such periodic alteration in segments' definitions could be impractical and will not result in robust value proposition for the Bank. Therefore, we propose an objective method of targeting prospect customers such as customer value (Net Present Value) based differentiation, where the NPVs (Net Present Value) are determined using an advanced predictive analytics method of behaviour segmentation. The benefits of portfolio segmentation to the credit marketer does exist, since there could be few segments with a higher value compared to few others with lower value.

2.0 OBJECTIVES

The broad objective of this paper includes providing a method of values -based differentiation of customers that helps create customer segments using an advanced data mining method and also to compare such incremental gains with the scenario of randomness (absence of any model, "BAU"). Thus, on a test data portfolio, we propose to establish that for few segments there exist positive and definite business benefits of implementing such a model, demonstrating that the segments vary in their profitability behaviour. The next section highlights the methodology of customer value differentiation.

3.0 METHODOLOGY & SEGMENTATION

The methodology of market segmentation is the division of markets into relatively homogeneous subgroups or target markets for the sake of broader strategy formulation and tactical decision making (Weinstein, 1993). Segmentation could also use statistical techniques to combine attitudinal and demographic data to develop groups that are easier to target (Berry and Linhoff, 1997; Christine, 1996 and Standford, 2000).

Therefore, the understanding of profitability in a banks' portfolio would typically depend upon a segmentation scheme separating borrower accounts into pools (segments) that are homogeneous with respect to customer behaviour. These pools (segments) then form the basis for estimating profitability measure within the segment where the customer behaviour is homogeneous. There could exist few challenges to the bank on the question of pooling. Lloyds Bank (2003) in their comments to Basel has also submitted these challenges. There is no generally accepted definition of homogeneity that forms an objective basis for determining appropriate segmentation. Kelly (2003) defines homogeneity as customers within a pool having a common default affinity. The basic problem is homogeneity over default affinity (or profit affinity) does not necessarily imply homogeneity over customer behaviour characteristics. Since, the profit (or default) is a resulting outcome, using profit levels (default rates) to create bands of profit (default grades) could result in biased methods and hence such objective segmentation methods cannot be used in models. Bag and Hosamane (2008) have suggested using non objective method of segmentations for retail assets based on risk characteristics. In view of the above limitation, instead of an objective segmentation (using Profit), a non objective segmentation method of natural grouping of data (natural clustering) using the risk characteristics of the borrowers may be an appropriate method of segmentation. Therefore, non objective segmentation method could imply homogeneity in customer behaviour.

An ideal profit measure is the net discounted cash accruals (NPV) from a customer segment adjusted against their net losses due to their expected defaults net of recoveries. This NPV at segment level is used for average NPV here. This comprises the K means nearest neighborhood clustering to create homogenous clusters of data which are heterogeneous across all clusters. The input to a k-means clustering would be the number of clusters which is refined iteratively. For natural grouping of data, clustering is used using two of the prominent techniques that includes; Hierarchal and K-Means. In case of larger number of attributes and bigger sample sizes or presence of continuous variables, K means is a superior method over Hierarchical. In our proposed K means model, character attributes were treated with suitable indicator transformations. Extreme Values of the numeric attributes were treated for outliers. Further, a multivariate collinearity check was also done for the variables within the model to eliminate any collinearity bias. The next section describes the method of revenue computation.

For each segment, we extend the approach of Perli and Nayda (2004) to measure the revenue due to a given outstanding balance for a given account at time t. Revenue on portfolio segments is due to the price income (interest income) and non price income (fees). Expenses are incurred on accumulated losses, funds and operating and marketing expenses including collection expenses. We simplify Peril and Nayda (2001) to assume that non-interest income is a constant fraction of outstanding balances. We also assume that, at default, the losses are the outstanding balance net of the recovered amount. Hence a constant fraction of borrowers in each segment pay the annual fee, the late fee, etc. Total revenue, is therefore, the sum total of all price and non price income. Similarly, total expenses are the sum of interest and non interest expenses. Interest expenses are the cost of funds and non interest expenses are the operating, marketing, etc. Interest expense is also related to the outstanding balances, since a portion of that needs to be financed by the bank. We assume that non-interest expenses are incurred on a per-account basis and, therefore, are a constant percentage, of outstanding balances.

Table 1 provides the parameters for calculating the net revenue for each of the segments. For the purpose of simplicity, the above parameters are assumed equal across all segments. However, in practice, these parameters are bound to vary across segments. Non-interest income factor, non interest expense factor, recovery factor, loss given default factor, etc are all assumed to be constant across all the groups. We calibrate our cluster model on the test portfolio which results in four significant groups. The results of the model are described in the next section.

4.0 MODEL RESULTS

As shown in the Chart 1 the implementation of the segmentation model on our test data of 30,000 accounts results in the distribution of four segments is robust groups of C (22%), D (29%), B (14%) and A (35%). The size of each segment is a good representation from the population drawn.

Table 1: Revenue Computation Parameters

r	average annual interest rate over balance	34%
λ	average fraction of non interest income over opening balance	2%
γ	loss given default (1- recovery ratio)	50%
cof	cost of funds (annual interest expenses)	15%
ψ	average fraction of non interest expenses over opening balance	15%

Source: Sample of Revolving Retail Assets RBI, June 2005

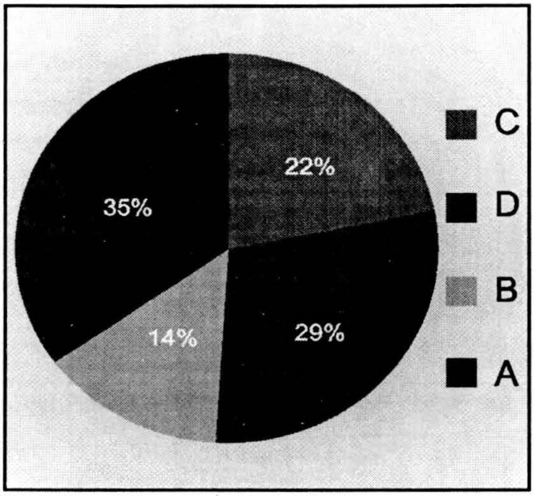

Chart 1: Size of Various segments

In the BAU (Business As Usual) scenario, the average profitability from the target base is lower. As shown in Table 2, the gain in our model scenario over a random BAU scenario stands at Rs. 168 and Rs.29 per account for the top two segments respectively. The average NPV for a segment is the ratio of the present value of net cash flows for the segment to the number of accounts in that segment. The BAU profitability is the average NPV calculated as the portfolio total to the total number of accounts (30,000). From the K means model, the segmentation drivers (characteristics) for each of the four groups include, Home Ownership, Occupation, Age of the Relationship, Balance Amount, Fees and Payment Amount, which includes a balanced mix of transaction, delinquency and demographic attributes. Table 3 provides the segments profile on the general characteristics of the customers. A closer look at the profile amongst the segments also reflects behavioural differences across the segments.

Thus it appears that the four proposed segments are heterogeneous across other groups and homogenous within each group. The next step involves fitting a Balance Amount model to be able to obtain projected revenue in future time period, shown in Table 4.

Table 2: NPV Summary

Segment	%age (Volume) i	NPV ('000 Rs.) ii	NPV/ Account (Rs.) iii (=ii/i)	?NPV over BAU (Rs.) iv (=iii-209)
C	22 (6,600)	1569	238	29
D	29 (8,700)	1530	176	-33
B	14 (4,200)	1585	377	168
A	35 (10,500)	1590	151	-58
ALL	100 (30,000)	6274	209	0

Source: Sample of Revolving Retail Assets for an Indian Bank, Reserve Bank of India, June 2005

Table 3: Segment Means Profile

Segment	Customer Age (years)	All Trade Balance (Rs.)	Annual Income (Rs.)	Outstanding Balance (Rs.)	Savings Balance (Rs.)	Delinquent Amount (Rs.)
C	27	27,759	3,33,268	22,207	1,66,556	262
D	46	1,31,787	3,66,630	1,83,315	36,66,300	0
B	29	66,229	7,94,749	1,44,607	3,97,375	604
A	52	47,206	8,04,428	4,02,214	8,04,4280	0

Source: Sample of Revolving Retail Assets for an Indian Bank, Reserve Bank of India, June 2005

Table 4: Forecast Model Results

Model Dependent Variable = Balance Amount (B_t)				
Model Variables	**Segment A**		**Segment B**	
	Estimate	**t Value ***	**Estimate t**	**Value ***
Previous Balance Amount (B_{t-1})	0.3	130	0.2	89
Vintage (MOB_t)	2.8	25	3.16	20
Delinquency Period ($Delq_t$)	6.03	9	7.41	6
	Segment C		Segment D	
Previous Balance Amount (B_{t-1})	0.1	109	0.3	115
Vintage (MOB_t)	3.05	23	3.2	25
Delinquency Period ($Delq_t$)	7.1	8	6.54	7

Source: Sample of Revolving Retail Assets for an Indian Bank, Reserve Bank of India, June 2005

The four regression models provide the forecast model values of Balance for the future time period during the life of the product. These models choose previous balance amount, vintage and delinquency period as the determinants of balance amount (B_t). The slopes of these variables are different for the four segments. Finally, the parameters mentioned in Table 1 are used to arrive at the cash flows for all segments by multiplying with the projected estimates of outstanding balance from the model in Table 4.

5.0 CONCLUSIONS

We presented a simple implementable credit marketing strategy that is implemented. These can have critical business implications. The bank finally has a informed picture of customer prospects for prospect targeting campaigns. It actually provides significant business insights to the bank. The banker has the flexibility to compare the NPV across products and also across customers. There could exist multiple applications of this analysis, with account level profitability established such as the portfolio actions on specific account management decisions.

There exists few gaps in these study and therefore may be highlighted for future research. Beyond profitability targeting approach needs to encompass optimization on other decision parameters such as market acceptance, etc. Larger Test Data that also includes both customer transaction information could provide greater accuracy and timeliness of the NPV analysis. Such analysis should be extended across all other products within the bank to understand the significance of generic segmentation vs product specific segmentation. Last but not the least, the incremental NPV gains should be validated by comparing with the implemented results of a campaign over time to close the loop of a full cycle model implementation.

REFERENCES

Art, Weinstein, et al. (1993): Hand Book of Market Segmentation; Mc Graw Hill.

Bag, D. and Hosamane, D. M. (2008). "Retail Assets Management in Indian Banking: Comparison of an Internal Model over Basel", International Journal of Rural Development & management Systems, December, Vol1, No.1.

Berry, M. and Linhoff, G. (1997). "Data Mining Techniques", John Willey.

Christine, E. (1996). "Marketing of Financial Services", Butterworth Heinemann.

Drumond, G. and Evison J. (2001). "Strategic Marketing", The Chartered Institute of Marketing, B & H.

Kelly, Shanon (2003): "Using Cluster Analysis for Retail Portfolio Segmentation in an Economic Capital Model:' Homogeneity by Common Default Behaviour over Time", NESUG03, Federal Reserve Bank of Philadelphia.

Lyods Bank (2003): "Comments Relating to Segmentation of Retail Credit Portfolios" http://www.bis.org/bcbs/cp3/allandsha.pdf); April 2003

Perli, R. and Nayda, W. (2004) "Economic and Regulatory Capital Allocation for Revolving Retail Customers", Journal of Banking & Finance, Elsevier, vol. 28(4), pages 789-809.

Stanford, G. (2000). "Data Mining in the Insurance Industry", SAS Inc.

Tina, H. (2000). "Financial Services Marketing", Prentice Hall, London.