Lifetime Value in the Bank Industry

Problem submitted by the National Bank of Canada

Sixth Montreal Industrial Problem Solving Workshop

August 17–21, 2015

General Context

- Acquiring and retaining profitable customers is an ever-growing challenge for banks.
- Customer demographics, buying behavior and needs are evolving rapidly.
- The competition is aggressive and adapting rapidly.
- Banks now need a 360-degree view of each customer in order to focus their resources efficiently.

National Bank of Canada

- These challenges are pushing the National Bank of Canada to gain a better understanding of the drivers behind its Customer Lifetime Value (CLV).
- The Client Intelligence & Modelization Team has been asked to develop methods to evaluate the CLV.
- The CLV will help the National Bank define, understand, and predict the relationship between a given client and the Bank.

The Problem

- The CLV is a simple concept in itself, but difficult to implement in a complex business context.

- A very large amount of data must be taken into account, for example:
  - Notions of client acquisition and attrition;
  - Holding of diverse banking products and services, their volume, usage, and profitability;
  - Other clients' characteristics such as geographical, demographic, and market data.

The Problem

- This data often needs a lot of cleaning and manipulation to be usable and meaningful.
- The profiles, products, or services of the Bank's clients vary greatly; so do their behaviours and expectations.
- The team's goal for this workshop was to explore and test different approaches to compute the Customer Lifetime Value.

Data Description: Commercial Data

- Aggregated monthly data from 2013 to 2015.
- "VECT_IN" binary vector of length 24, which stands for the products: savings (3), loans (5), transactions (16). Contains 1479 distinct vectors.

VECT_IN	NB_VECT_IN	VALUE
000001011010000001100000	27240	189.2424
000001011010000001110000	21855	204.9733

Data Description: Retail Data

- Original data: 4831 retail clients (rows), 6130 attributes (columns).
- Cleaned data: 4831 rows and 88 columns.
- Data for a particular month.
- Main attributes are:
  - Demographic attributes (age, gender, marital status, ...).
  - Product information (tenure, holding, balance).
  - EFT (Electronic Fund Transfers), AML (Anti-Monetary Laundering), Aggregated fees.

More values for each product could improve the reliability of the suggested solution.

Relevant Literature

1. We focused on exploring an indicator called **Customer Lifetime Value (CLV)**.
2. There exists several studies focusing on CLV in the retail banking industry.
3. The approaches of [3], [2], and [1] informed our analysis and approach to modelling the CLV.
4. We want a model or algorithm for estimating and predicting the CLV in order to drive marketing decisions and customer relationship management.

Relevant Literature

1. [3] proposed a customer valuation model based on a combination of Markov chains and classification and regression trees (CART).
2. Use German bank data.
3. Focus on parsimonious model based on homogeneous customer segments.
4. Four main factors: age, demographics and lifestyle, type and intensity of product usage, and activity level.
5. Factors measured by multiple indicators used as predictor variables for profit contribution.

Relevant Literature

1. CART analysis to cluster the client base into several homogeneous sub-groups.
2. Each homogeneous sub-group was considered a state of a Markov chain.
3. Moves between states over time with transition probabilities.
4. Probabilities estimated by counting the number of customers who moved between two states and dividing by the total number of customers.

Relevant Literature

1. [2] uses CLV for the modelling and prediction of a certain type of customer behaviour.
2. Denote the usage of a product j, during time period t, by customer i by the variable $x_{i,j,t}$.
3. CLV for customer i from time t to $t+h$ is defined as

$$CLV_{i,t} = \sum_{k=1}^{h} \sum_{j=1}^{q} \frac{1}{(1+r)^k} CF_{i,j,t+k}. \tag{1}$$

Relevant Literature

1. $CF_{i,j,t}$ denotes the net cash flows yielded by the transaction on product j.
2. $r > 0$ is a relevant discount rate for the time period $[t, t+1)$.
3. π_j, for the j-th product during the time period, denotes the average marginal profit per unit of product usage.
4. $CF_{i,j,t} = \pi_j x_{i,j,t}$

Relevant Literature

1. [2] performs a classification of customers.
2. Goal of avoiding misclassification based on a CLV-sensitive loss function.
3. Methods (including logistic regression, decision trees, and neural networks) were used and compared for the classification procedure.

Objectives

Aim: Develop a model for Client Lifetime Value

Objectives:

1. Identify and segment relevant variables.
2. Develop CLV model.
3. Develop a predictive model to help determine future marketing strategies based on CLV.

First objective

Identify and segment relevant variables

- Volume of data
- Understand drivers of CLV

Approaches used

- Generalized linear models (GLM)
- Stepwise regression
- Classification and regression trees (CART)
- Principal components analysis (PCA)
- Self-organizing maps (SOM)

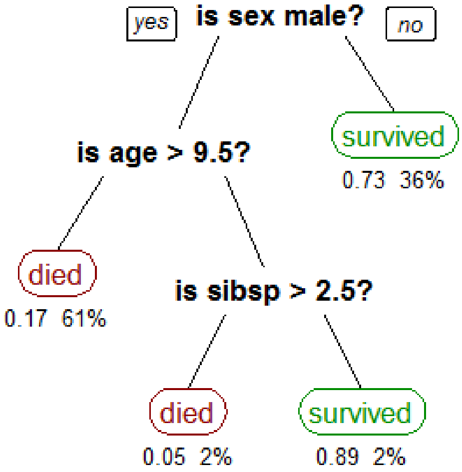
GLM

Express CLV_i as a function of

- Products held: 0-1 vector, e.g., $(1, 0, 1, 1, 0, 1)$;
- Demographics: age, gender, income, zip code region;
- Client profile and behavior: credit score, average balance in checking account, number of transactions.

$$CLV_i = f(\beta_{0i} + \beta_{1i}\xi_{1i} + \beta_{2i}\xi_{2i} + \ldots + \epsilon_i)$$

Classification and regression trees (CART)



PCA

- Dimension reduction

The vector of explanatory variables

$$(\xi_1, \xi_2, \xi_3, \ldots, \xi_n)$$

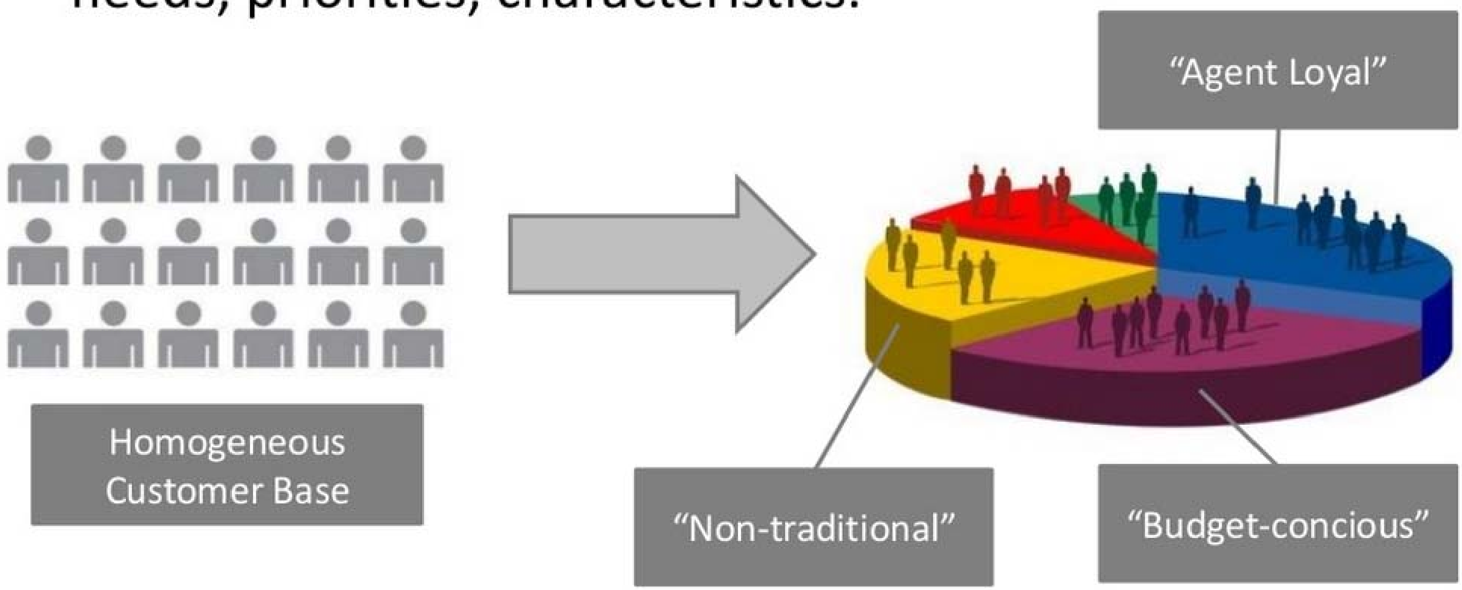
is mapped into

$$\left(\sum_{i=1}^{n} w_i^{(1)} \xi_i, \sum_{i=1}^{n} w_i^{(2)} \xi_i, \ldots, \sum_{i=1}^{n} w_i^{(n)} \xi_i \right).$$

- A few components explain the variability in the data
  - Line of credit
  - Personal loans (e.g., car loan, student loan)
  - Mastercard

Customer Segmentation

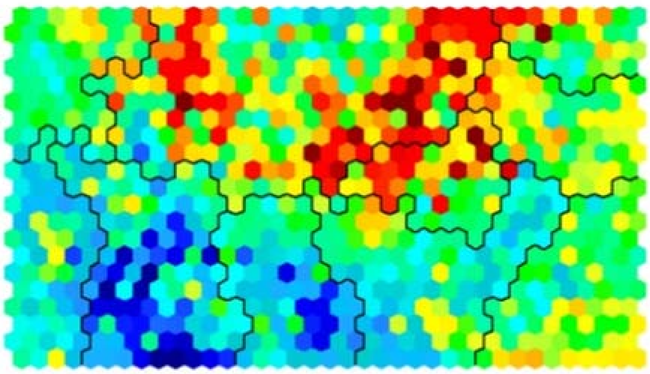
- Customer segmentation is the application of clustering techniques to customer data
- Identify cohorts of “similar” customers – common needs, priorities, characteristics.



Self-Organising Maps

A Self-Organising Map (SOM) is a form of unsupervised neural network that produces a low (typically two) dimensional representation of the input space of the set of training samples.

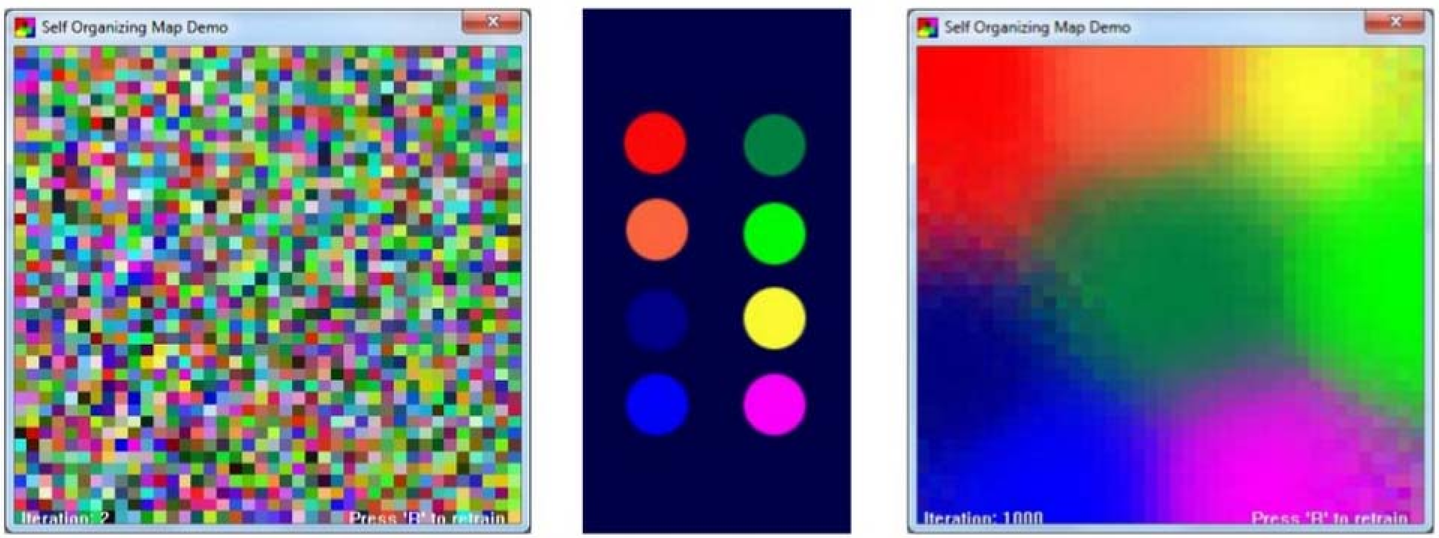
- First described by Teuvo Kohonen (1982) ("Kohonen Map")
- Over 10k citations referencing SOMs – most cited Finnish scientist.
- Multi-dimensional input data is represented by a 2-D "map" of nodes
- Topological properties of the input space are maintained in map



Self-Organising Maps

Example – Color Classification

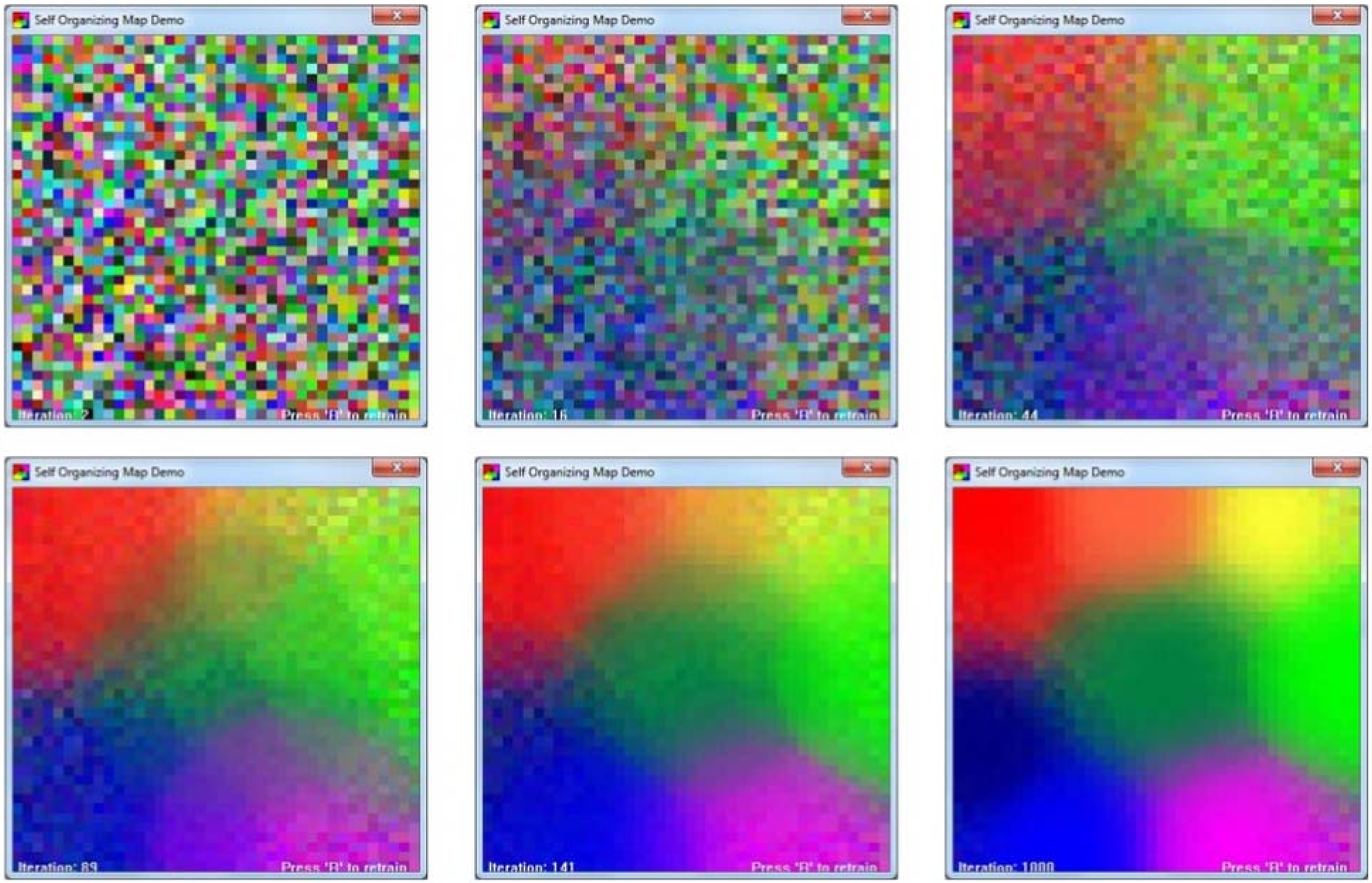
- SOM training on RGB values. (R,G,B) (255,0,0)
- 3-D dataset -> 2-D SOM representation
- Similar colours have similar RGB values / similar position



Think of different colors as different customers having different set of attributes (R, G, B's)

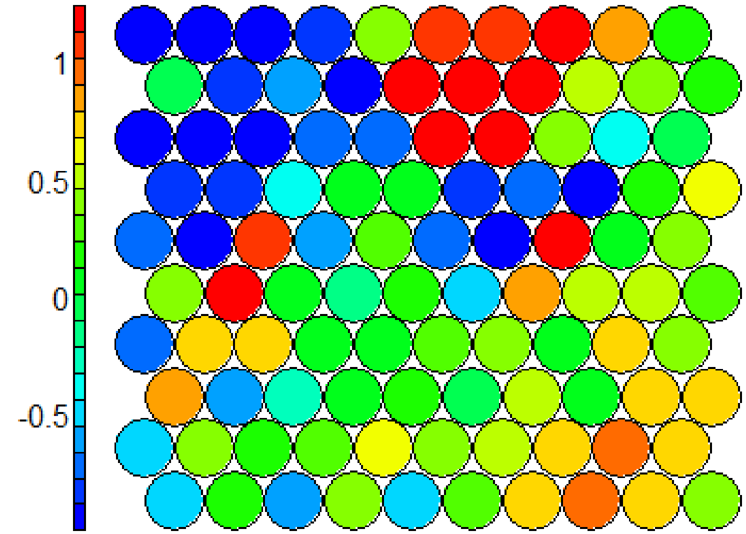
Self-Organising Maps

SOM Colors



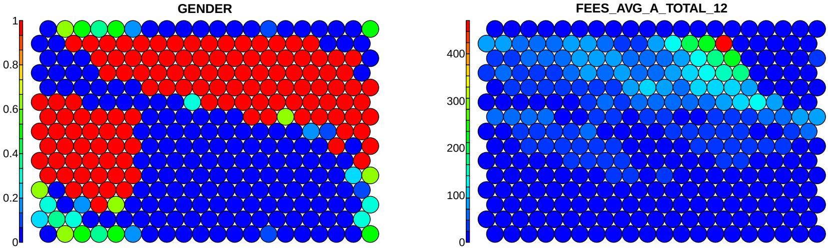
SOM →
extract clusters →
reduce dimensionality →
simpler Markov chains
with fewer transition
probabilities.

SOM HeatMap

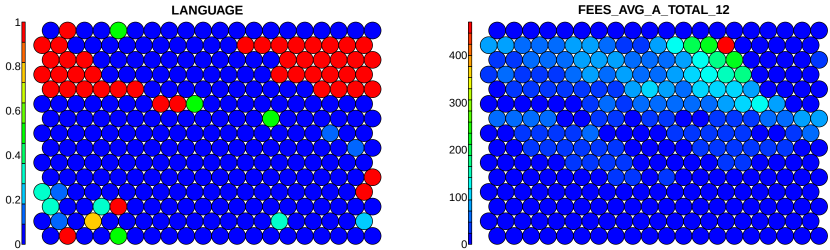


The output clustering of the SOM, **retail banking** information is cash flow, demographics and produc holdings

Gender vs. Fees_Avg_Total_12



Language vs. Fees_Avg_Total_12



Second objective

Develop CLV model

- Literature review
- Static: GLM
- Dynamic: modelling transition of states

Highlights

- Importance of segmentation
- (Semi-)Markov chain
- Logistic regression

Markov chains

Vector of product holdings $(x_{t,1}, \ldots, x_{t,q})$

- Example: (1,0,1,1,0,1)
- q products: 2^q possible vectors

Importance of segmentation

$$P = \begin{pmatrix} P_{1,1} & P_{1,2} & \cdots & P_{1,q} \\ P_{2,1} & P_{2,2} & \cdots & P_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ P_{q,1} & P_{q,2} & \cdots & P_{q,q} \end{pmatrix}$$

Markov chains

- Segmentation of transitions by client characteristics.
- Clients spend a lot of time in a given state.

A Markov chain implies a geometric distribution for the amount of time spent in a given state.

$$\Pr(T = n) = p^{n-1}(1-p), \quad n = 1, 2, \ldots$$

- Modelling constraint

Markov chains

- Possible solution: model separately the time spent in a state and the transitions.
- Semi-Markov chain

$$P = \begin{pmatrix} x & P_{1,2} & \cdots & P_{1,q} \\ P_{2,1} & x & \cdots & P_{2,q} \\ \vdots & \vdots & \ddots & \vdots \\ P_{q,1} & P_{q,2} & \cdots & x \end{pmatrix}$$

Logistic regression

To model the transition probabilities as a function of explanatory variables, we can use a logistic regression.

$$a_{j,k} = \beta_{0,jk} + \beta_{1,jk}\xi_1 + \beta_{2,jk}\xi_2 + \dots$$

$$\Pr(\mathbf{x}_{t+1} = k \mid \mathbf{x}_t = j, \xi) = \frac{\exp(a_{j,k})}{\sum_{j=1}^{q} \exp(a_{j,k})}$$

Third objective

Develop a predictive model to help determine future marketing strategies based on CLV

Thank you!

References

[1] Y. Ekinci, F. Ülengin, N. Uray, and B. Ülengin. Analysis of customer lifetime value and marketing expenditure decisions through a Markovian-based model. *European Journal of Operational Research*, 237:278–288, 2014.

[2] N. Glady, B. Baesens, and C. Croux. Modeling churn using customer lifetime value. *European Journal of Operational Research*, 197:402–411, 2009.

[3] M. Haenlein, A. M. Kaplin, and A. J. Beeser. A model to determine customer lifetime value in a retail banking context. *European Management Journal*, 25(3):221–234, 2007.